

# **Identification of Genes Regulating Invasive Behaviour in a Model of Human Cancer**

**A thesis submitted for the award of the degree of Doctor  
of Philosophy**

**By**

**Azeem Arshad**



**Department of Pathology**

**School of Medicine, Cardiff University**

**September 2010**



# **Identification of Genes Regulating Invasive Behaviour in a Model of Human Cancer**

**A thesis submitted for the award of the degree of Doctor  
of Philosophy**

**By**

**Azeem Arshad**



**Department of Pathology**

**School of Medicine, Cardiff University**

**September 2010**

UMI Number: U517256

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U517256

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346



# DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed Aze M. Q. (candidate) Date 8/3/2011

## STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed Aze M. Q. (candidate) Date 8/3/2011

## STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references.

Signed Aze M. Q. (candidate) Date 8/3/2011

## STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed Aze M. Q. (candidate) Date 8/3/2011

## STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans after expiry of a bar on access previously approved by the Graduate Development Committee.

Signed ..... (candidate) Date .....

# Acknowledgements

I first and foremost wish to thank my supervisor Dr. Zuhuri Poghosyan. Without her direction, support this thesis would not be possible. I hope this work will form the start of future collaborations.

I am indebted to Dr. Rachel Errington who provided me with vital guidance in the study of cell motility. This thesis would not exist if it wasn't for The CBS department at Cardiff University. The work of Dr. Peter Giles, Megan Musson and Dr. Claudia Consoli was vital to the Ph.D thesis.

A special thank you goes to Dr. Martyn Brown, who provided me with novel methodology to study cell motility.

I would like to give many thanks to my colleagues in the Pathology Department. This thesis would not be possible without combination of the support and expertise

I would like to offer a special thank you to my close friends, Chris Ward, Kee Tang and Simon Wong. You guys are legends! I could not have done this Ph.D without you. I would also like to thank Jack Turner, Heather Richardson and Duncan Meacher and the rest of the Cardiff University Mountaineering Club. All you guys kept me sane!

I would like to acknowledge the financial support by Cardiff University and Cancer Research UK.

I would like to extend my gratitude to my family for their continuous support.

Finally I would like to extend a very special thanks to my darling wife, Nafeesa. Without your tireless and amazing support I would not be in this position of such success.

## Abstract

For this project I aimed to identify and investigate genes which regulate an invasive phenotype in thyroid carcinomas. To investigate this aim, I utilised an established *in vitro* model system which involves the introduction of the oncogenes RASV12, BRAF<sup>V600E</sup> and RET/PTC1 by retroviral infection into normal primary thyrocytes to represent the *in vivo* tumour situation. I adopted Affymetrix microarray gene expression profiling of the oncogene infected primary thyrocytes to identify candidate genetic markers facilitating thyroid invasion. In parallel I performed a literature analysis to identify additional markers of invasion. The genes identified from this analysis were Slug and Osteopontin. Osteopontin is over-expressed and enhances papillary carcinoma invasion in a Fisher rat cell line model of thyroid carcinoma. In tissue samples, Slug is over-expressed in papillary carcinomas. Currently, there is limited research on how Slug and Osteopontin regulate thyroid tumour invasion. I have investigated the role of Slug and Osteopontin in tumour invasion through the employment of RNA interference. I have analysed the role of Slug and Osteopontin in a number of invasive mechanisms. The mechanisms analysed included degradation of the extracellular matrix, cell motility, regulation of the actin cytoskeleton during cell migration, cell-matrix interactions and cell-cell interactions. The results indicate that Osteopontin influences the actin cytoskeleton by possibly regulating the formation of contractile stress fibres required for forward movement. Slug mediates tumour invasion by deregulating the cell-cell interactions which are important for maintaining normal epithelial architecture.

# Contents

Declaration.....	I
Acknowledgements.....	II
Abbreviations.....	III
Abstract.....	IV
Contents.....	V
List of Figures.....	VI
List of Tables.....	VII

## **PART I**

<b>1.0 Background.....</b>	<b>1</b>
<b>1.1 The Thyroid Gland.....</b>	<b>2</b>
<b>1.2 Thyroid Neoplasia.....</b>	<b>4</b>
1.2.1 Follicular Adenoma (FA).....	4
1.2.2 Follicular Carcinoma (FTC).....	4
1.2.3 Papillary Carcinoma (PTC).....	5
1.2.4 Poorly differentiated and undifferentiated carcinomas.....	6
<b>1.3 Molecular events involved with thyroid cancer.....</b>	<b>7</b>
1.3.1 Mitogen-activated protein kinase (MAPK) signalling pathway.....	7
1.3.1.1 RET.....	8
1.3.1.2 Ras.....	11
1.3.1.3 RAF family.....	14
<b>1.4 Gene expression profiling of thyroid lesions.....</b>	<b>19</b>
1.4.1 Differentiating Thyroid tumours .....	19
1.4.2 Molecular characterisation of Thyroid Tumours.....	23

1.4.3 Genetic Profiling of Thyroid Tumour cell lines .....	31
1.4.4 Overall perspectives.....	32
<b>1.5 In vitro models of thyroid neoplasia.....</b>	<b>34</b>
<b>1.6 Aims, Objectives and Experimental strategies.....</b>	<b>35</b>
<b>2.0 Materials and Methods.....</b>	<b>38</b>
<b>2.1 Generation of oncogene expressing thyroid colonies.....</b>	<b>39</b>
2.2.1 Cell culture .....	39
2.2.2 Preparation of viral supernatants.....	40
2.2.2.1 Preparation of Retroviral supernatant.....	40
2.2.3 Viral infections.....	41
<b>2.3 Total RNA extraction.....</b>	<b>42</b>
<b>2.4 Sequencing of B-Raf.....</b>	<b>43</b>
<b>2.5 One-cycle eukaryotic labeling of gene chips.....</b>	<b>44</b>
2.5.1 Sample preparation.....	44
2.5.2 Quantification of RNA.....	45
2.5.3 One-cycle cDNA synthesis.....	45
2.5.4 First-strand cDNA synthesis.....	46
2.5.5 Second-strand cDNA synthesis.....	47
2.5.6 Cleanup of Double-Stranded cDNA.....	48
2.5.7 Synthesis of Biotin-labelled cRNA.....	49
2.5.8 Cleanup and quantification of Biotin-Labelled cRNA.....	50
2.5.9 Fragmentation of the cRNA for Target Preparation.....	50
2.5.10 Eukaryotic Target Hybridisation.....	51
2.5.11 Washing, staining and scanning the array.....	52
<b>2.6 Invasion Assay.....</b>	<b>53</b>



<b>3.0 Results</b>	<b>55</b>
<b>3.1 The status of BRAF<sup>V600E</sup> mutation in thyroid cell lines</b>	<b>56</b>
3.1.1 Objectives	57
3.1.2 Multiple BRAF isoforms are detected in thyroid cell lines	57
3.1.3 The BRAF <sup>V600E</sup> mutation is detected in the BCPAP and K1 papillary carcinoma cell lines	59
3.1.4 Summary	62
<b>3.2 Microarray Expression Gene Profiling</b>	<b>65</b>
3.2.1 Objectives	65
3.2.2 Preparation of oncogene infected thyrocytes	66
3.2.3 Microarray quality control metrics	66
3.2.3.1 Spatial Artifacts	67
3.2.3.2 RNA degradation	68
3.2.3.3 Percent present genes and internal control genes	70
3.2.3.4 QC Summary	72
3.2.4 Microarray Analysis Design	73
3.2.5 Gene Profiling Analyses	73
3.2.5.1 Statistical t-test analysis of RASV12, BRAF <sup>V600E</sup> and RET/PTC1	74
3.2.5.2 Analysis of Variance (ANOVA) analysis	80
3.2.6 Invasive behaviour of the in vitro models	84
3.2.7 Identification of candidate genes	86
3.2.8 Summary	95

## **PART II**

<b>4.0 Background</b>	<b>103</b>
<b>4.1 The SNAIL family of transcriptional repressors</b>	<b>104</b>

<b>4.2 Osteopontin (OPN)</b>	<b>108</b>
4.2.1 <i>Structure of OPN</i>	109
4.2.2 <i>OPN receptors</i>	112
4.2.2.1 <i>CD44</i>	112
4.2.2.2 <i><math>\alpha\beta</math>Integrins</i>	116
4.2.3 <i>Regulation of OPN expression</i>	117
4.2.3.1 <i>Ras</i>	117
4.2.3.2 <i>Rous sarcoma oncogene (Src)</i>	117
4.2.3.3 <i>Tcf-4</i>	118
4.2.3.4 <i>AP-1</i>	118
4.2.4 <i>OPN and proteinase activity</i>	119
4.2.5 <i>The Hepatocyte Growth Factor (HGF)-cMet system</i>	120
4.2.6 <i>OPN in thyroid cancer</i>	121
<b>4.3 Tumour Invasion and Metastasis</b>	<b>122</b>
4.3.1 <i>Epithelial to mesenchymal transition (EMT)</i>	123
4.3.2 <i>Cell adhesion and cytoskeletal rearrangement</i>	124
4.3.2.1 <i>Cell Adhesion</i>	124
A) <i>Integrins</i>	124
B) <i>Cell-matrix interactions</i>	127
C) <i>Actin filaments forming the basis of cell movement</i>	132
D) <i>Cell-cell interactions</i>	139
4.3.2.3 <i>Proteolysis</i>	142
A) <i>Matrix Metalloproteinases</i>	142
B) <i>Urokinase Plasminogen Activator System</i>	144
<b>4.4 Aims, Objectives and Experimental Strategy</b>	<b>146</b>
<b>5.0 Materials and Methods</b>	<b>150</b>
5.1 <i>Cell culture</i>	151
5.2 <i>Preparation of Lentiviral supernatants</i>	151

<b>5.3 Cell Lysis</b>	<b>155</b>
<b>5.4 Protein Quantification</b>	<b>155</b>
<b>5.5 Sodium Dodecyl Sulfate (SDS) Polyacrylamide Gel Electrophoresis (PAGE) Western Immunoblotting</b>	<b>156</b>
<b>5.6 Zymogram Analysis</b>	<b>157</b>
<b>5.7 OPN Enzyme-linked immunosorbent assay (ELISA)</b>	<b>158</b>
<b>5.8 Immunofluorescence</b>	<b>159</b>
<b>5.9 Phalloidin Staining</b>	<b>159</b>
<b>5.10 Mini-preparation of plasmid DNA</b>	<b>161</b>
<b>5.11 cDNA synthesis and Reverse Transcription Polymerase Chain Reaction (RT-PCR)</b>	<b>163</b>
<b>6.0 Results</b>	<b>166</b>
<b>6.1 The role of Slug and OPN in Thyroid Tumour Invasion</b>	<b>167</b>
<b>6.2 Characterising the invasive behaviour across a panel of thyroid cancer cell lines</b>	<b>169</b>
<b>6.3 Characterising the role of Slug and OPN in tumour invasion</b>	<b>173</b>
<i>6.3.1 Characterisation of Slug in thyroid cell lines</i>	<i>174</i>
<i>6.3.2 Characterisation of OPN in thyroid cell lines</i>	<i>178</i>
<b>6.4 The shRNA mediated knockdown of Slug and OPN</b>	<b>180</b>
<b>6.5 Summary</b>	<b>188</b>

<b>6.6 Invasive processes mediated by Slug and OPN</b>	<b>192</b>
6.6.1 <i>Proteinase activity mediated by Slug and OPN</i>	192
6.6.1.1 <i>Summary</i>	195
6.6.2 <i>Cell migration</i>	198
6.6.2.1 <i>Analysis of the actin cytoskeleton during cell migration</i>	198
6.6.2.2 <i>Analysis of cell motility</i>	202
6.6.2.2.1 <i>The effect of Slug and OPN down-regulation on cell growth</i>	203
6.6.2.2.2 <i>Random walk analysis</i>	206
6.6.2.2.3 <i>Analysis of velocity during cell migration</i>	208
6.6.2.3 <i>Summary</i>	211
<b>6.7 Induction of Mesenchymal to Epithelial Transition by Slug and OPN</b>	<b>213</b>
6.7.1 <i>Analysis of WAVE and N-WASP during cell migration</i>	214
6.7.2 <i>Summary</i>	217
<b>6.8 The role of HGF in Slug and OPN-mediated cell migration</b>	<b>219</b>
6.8.1 <i>Assessment of CD44v and the implication for HGF-mediated invasion</i>	223
6.8.2 <i>Summary</i>	226
<b>6.9 Analysis of focal adhesion formation during cell migration</b>	<b>228</b>
6.9.1 <i>Analysis of Vinculin and Cortactin</i>	228
6.9.2 <i>The regulation of Integrin <math>\beta</math>3 localisation</i>	231
6.9.3 <i>Summary</i>	233
<b>6.10 Cell-cell adhesion mediated by Slug and OPN</b>	<b>236</b>
6.10.1 <i>Tight Junctions</i>	236
6.10.2 <i>Adherence junctions</i>	238
6.10.3 <i>Summary</i>	243
<b>7.0 General Discussion and Future work</b>	<b>246</b>
<b>8.0 References</b>	<b>259</b>





# List of Figures

## PART I

### Chapter 1.0 - Background

Figure 1. Representation of the thyroid gland detailing localisation and cells

Figure 2. Normal RET and oncogenic RET/PTC1

Figure 3. Ras activation and signalling

Figure 4. Schematic representation of B-Raf isoforms

Figure 5. Schematic diagram of the RAF family of proteins

### Chapter 3.0 - Results

Figure 6. B-Raf isoforms identified by Western blot.

Figure 7. Methodology of RT-PCR sequencing for the full length and shorter isoforms of B-raf.

Figure 8. Sequencing data of selected cell lines observing BRAF<sup>V600E</sup> mutation.

Figure 9. Spatial artifact quality control analysis.

Figure 10. RNA Degradation plots of RASV12, BRAF<sup>V600E</sup> and RET/PTC1 genechips.

Figure 11. Percent present genes and internal control genes analysis of RASV12, BRAF<sup>V600E</sup> and RET/PTC1 genechips.

Figure 12. Representative view and the number of invaded cells of the RASV12, BRAF<sup>V600E</sup> and RET/PTC1 models.

Figure 13. Expression profiles of Matrix Metalloproteinases determined in MADRAS.

Figure 14. Expression analysis of Profilin-1 using MADRAS

Figure 15. Expression analysis of S100A4 using MADRAS.

Figure 16. Expression analysis of Tenascin-C using MADRAS.

Figure 17. Expression analysis of Slug using MADRAS.

Figure 18. Expression analysis of OPN using MADRAS.

## **PART II**

### **Chapter 4.0 - Background**

Figure 19. A. Genomic structure highlighting OPN spliced isoforms.

Figure 20. Exon map of CD44 displaying spliced exons and conserved exons.

Figure 21. Schematic diagram of integrin activation.

Figure 22. Overview of focal adhesion formation and maturation.

Figure 23. Overview of RhoGTP family signalling during cytoskeletal rearrangement.

Figure 24. Overview of cell movement on a 2D substrate.

Figure 25. The Urokinase Plasminogen Activator System.

### **Chapter 5.0 – Results**

Figure 26. The invasive potential of cell lines derived from different thyroid pathologies using the Boyden chamber system.

Figure 27. Characterisation of Slug in thyroid cell lines

Figure 28. Immunofluorescent analysis of Slug expression in cell lines derived from various thyroid pathologies

Figure 29. Characterisation of Osteopontin in thyroid cell lines.

Figure 30. Confirmation of shRNA mediated knockdown of Slug.

Figure 31. Confirmation of shRNA mediated knockdown of OPN.

Figure 32. The invasive potential of the Slug Sh cell line.

Figure 33. The invasive potential of the OPN Sh cell line.

Figure 34. Analysis of proteinase activity in Slug Sh and OPN Sh.

Figure. 35. Immunofluorescent timecourse observing wound closure with F-actin staining. Figure 36. Growth curve to measure the number of K1, Slug Sh and OPN Sh cells every 5 hours.

Figure 37. Inter-mitotic time Vs. inter-mitotic distance for each complete generation of K1, Slug Sh and OPN Sh cells.

Figure 38. Generational Random Walk analysis for K1, Slug Sh and OPN Sh.

Figure 39. Preliminary experiment displaying the average velocity for each generation of K1, Slug Sh and OPN Sh cells.

Figure 40. Immunofluorescent analysis of F-actin and WAVE1 expression in Slug Sh and OPN Sh cells.

Figure 41. Immunofluorescent analysis of F-actin and N-WASP expression in Slug Sh and OPN Sh cells.

Figure 42. Preliminary experiment displaying Inter-mitotic time Vs. inter-mitotic distance for K1, Slug Sh and OPN Sh cells in a wound closure experiment.

Figure 43. Average velocity for the 1<sup>st</sup> generation of K1, Slug Sh and OPN Sh cells in a wound closure experiment.

Figure 44. Analysis of CD44 variants in Slug Sh and OPN Sh cells.

Figure 45. Surface expression of CD44 (all isoforms) in Slug Sh and OPN Sh cells.

Figure 46. Co-immunostaining to detect vinculin and phosphorylated-cortactin expression in ORI-3, K1, Slug Sh and OPN Sh cells.

Figure 47. Integrin $\beta$ 3 localisation in ORI-3, K1, Slug Sh and OPN Sh cells.

Figure 48. Analysis of tight junctions by Slug Sh and OPN Sh.

Figure 49. Analysis of E-cadherin expression and mRNA content.

Figure 50. Immunofluorescent analysis of E-cadherin and  $\beta$ -catenin expression.

Figure 51. Pan-cadherin expression in ORI-3, K1, Slug Sh and OPN Sh cells.

# List of Tables

## **PART I**

### **Chapter 1.0 – Background**

Table 1.0. RET rearrangements with fusion partners.

### **Chapter 2.0 – Materials and Methods**

Table 2.0. Retroviral plasmids, producer cells and selection

Table 2.1. First-Strand Master Mix

Table 2.2. Second-Strand Master Mix

Table 2.3. IVT reaction set-up

Table 2.4. IVT labelling kit components

Table 2.5. Fragmentation Mix components

Table 2.6. Hybridisation Cocktail Mix

Table 2.7. Automated protocol for scanning and staining

### **Chapter 3.0 - Results**

Table 3.0. Thyroid cell lines with associated origin and mutational status.

Table 3.1. Details of hierarchical clustering of genotype-specific genes in RASV12, BRAF<sup>V600E</sup> and RET/PTC1 oncogene infected thyrocytes



## **PART II**

### **Chapter 5.0 – Materials and Methods**

Table 5.1. Cells with corresponding culture conditions, source and usage.

Table 5.2 Lentiviral plasmids with producer cells used and corresponding selection

Table 5.3. Antibodies used for Immunofluorescence (IF) and Western Blot analysis (WB).

Table 5.4. PCR program for primers ordered through MWG.

Table 5.5. PCR program for GAPDH primers ordered from Biomol.

Table 5.6. Primer sequences detailing product size, number of cycles and melting temperature.

### **Chapter 6.0 – Results**

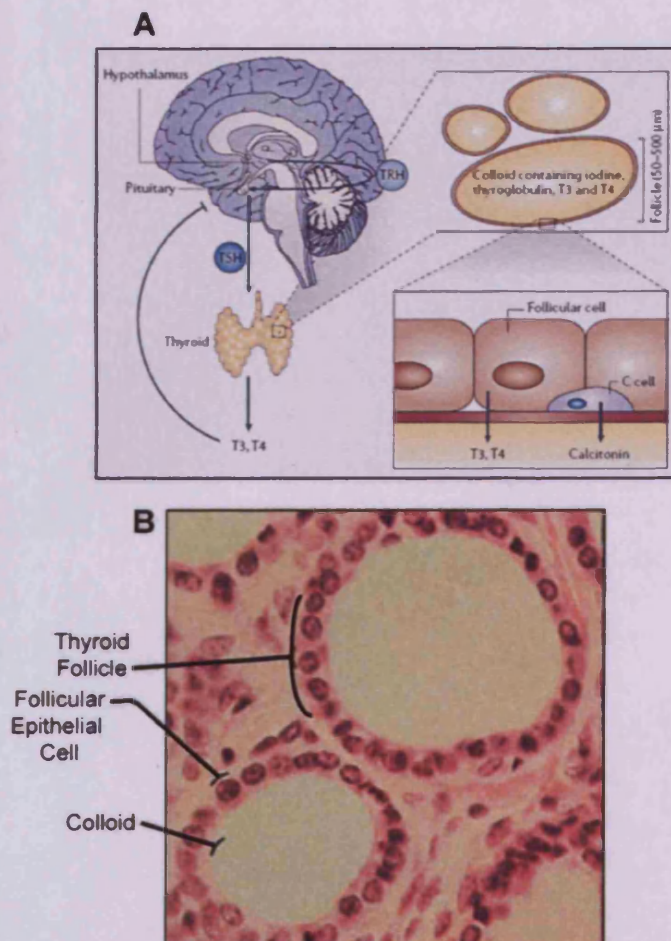
Table 6.0. The thyroid cell lines and their source used for the characterisation of invasive behaviour.

# **PART I**

## **1.0 Background**

## 1.1 The Thyroid Gland

The thyroid gland is located on the anterior surface of the trachea at the base of the neck. It is comprised of lobules which contain spherical follicles which are lined with follicular epithelial cells. The follicles contain colloid, a store for thyroglobulin which is a precursor of the thyroid hormones, L-triiodothyronine (T3) and L-thyroxine (T4). These hormones are regulated by the release of thyrotrophin-releasing hormone from the hypothalamus which induces the release of thyroid-stimulating-hormone (TSH) from the anterior pituitary gland. This stimulates the synthesis and secretion of T3 and T4 from the thyroid (Kondo et al. 2006) (Figure 1). Calcitonin producing C-cells are present at the junction of the upper and middle third of both thyroid lobes, and are situated within the follicles.



**Figure 1.** A, Representation of the thyroid gland detailing localisation and cells (Kondo et al. 2006). B, Histological section of a normal thyroid gland detailing follicles, epithelial cells and colloid (D'Alessandro, 2005).

Methods to investigate thyroid neoplasms prior to thyroidectomy include ultrasound imaging, nuclear scan and fine needle aspiration biopsy (FNA). The FNA offers the best preoperative diagnostic approach, with a reliability of 95% for the diagnosis of papillary thyroid carcinoma (PTC) (Zhao et al. 2008). This diagnostic approach is limited as it cannot differentiate follicular adenoma (FA) from follicular thyroid carcinoma (FTC) as these lesions have overlapping pathological features making a differential diagnosis difficult. Patients with indeterminate or suspicious tumours undergo thyroidectomy, and up to 80% of these particular surgeries are determined as FA (Nunez and Mendelsohn 1989). The sole criterion for the diagnosis of FTC is the demonstration of capsular or vascular invasion, ascertained by histological examination following thyroidectomy (Lewinski et al. 2000). There is a need for further improvements to diagnose and to identify differential genetic makers which may distinguish FTC from FA.

## 1.2 Thyroid Neoplasia

Approximately 90% of thyroid tumours are derived from follicular epithelial cells, whilst the rest are derived from other cell types (DeLellis 2006). These epithelial tumours are classified as either follicular adenomas, follicular or papillary carcinomas (differentiated thyroid carcinomas) and anaplastic carcinomas (undifferentiated thyroid carcinomas).

### 1.2.1 Follicular Adenoma (FA)

Follicular adenomas are benign tumours which are distinguished from the surrounding parenchyma by a 'well-defined' intact fibrous tissue capsule. The tissue capsule is a thin connective tissue layer comprising a mixture of adipose and connective tissues with a variable thickness (Mete et al. 2010). Thyroid FAs are associated with haemorrhage, cystic degeneration, ossification, calcification and fibrosis (Suster 2006).

### 1.2.2 Follicular Thyroid Carcinoma (FTC)

Defined by the World Health Organisation (WHO), follicular carcinomas are malignant epithelial tumours with evidence of follicular cell differentiation and comprise of 10-15% of all thyroid malignancies, with a higher incidence in women (DeLellis 2006). FTCs derive from FA, and are also encapsulated comprising of follicles or follicular cells arranged in a follicular, solid or trabecular pattern. FTCs displaying transcapsular and/or limited vascular invasion are classified as minimally invasive FTC. Those FTCs which display wide spread capsular and vascular



invasion are classified as widely invasive FTCs (Baloch and Livolsi 2002). The presence of limited capsular invasion is an indicative risk factor for metastasis and recurrence. The risk is increased in FTCs with limited capsular and vascular invasion, and these widely invasive FTCs are associated with a significant increase in mortality and morbidity (DeLellis 2006). Metastasis occurs through the haematopoietic system to secondary sites such as the brain, liver and lungs (Mazzaferri 1999). The presence of capsular and/or vascular invasion in FTC are the only histopathological differences which distinguish FTCs from FAs.

### *1.2.3 Papillary Thyroid Carcinoma (PTC)*

Papillary carcinomas are defined as malignant tumours which derive from follicular cells. They comprise ~90% of all thyroid malignancies, occurring mainly in women at a ratio of 4-5:1 (DeLellis 2006). The diagnostic features of PTC include characteristic papillae where the papillary cores are covered by neoplastic cells. They lack a tumour capsule and possess distinct histopathological features including nuclear enlargement and irregularity with finely dispersed chromatin, a highly folded nuclear membrane and an abnormal nuclear contour. Metastasis of PTC occurs through the lymphatic system (Al-Brahim and Asa 2006). Variants of classical PTC have been identified and these tumours present a diagnostic challenge as they may be encapsulated or possess nuclear features as seen with conventional PTC.

Papillary carcinomas may also present as 'microcarcinomas', and these are defined as papillary lesions measuring less than 1cm in size with gross extension beyond the capsule. This type of lesion is common, occurring in up to 24% of thyroidectomies unrelated to papillary carcinoma and in up to 30% of autopsies. Most well

differentiated tumours have good prognosis and with an increased survival rates as they are slow growing.

#### *1.2.4 Poorly differentiated and undifferentiated carcinomas*

Poorly differentiated tumours (insular thyroid carcinomas) are defined as follicular cell carcinomas with limited follicular cell differentiation. They occupy an intermediate position between differentiated carcinomas and undifferentiated thyroid carcinomas. Cytologically, these tumours have high cellular smears, scanty or no colloid. The cells appear singly and in syncytial with irregular aggregates. Nuclear overlapping and overcrowding, poorly formed micro-follicles, a high nuclear to cytoplasmic ratio, necrosis are common features (Kondo et al. 2006). Undifferentiated anaplastic thyroid carcinomas (ATC) are rare tumours occurring primarily in individuals above the age of 60. Typically, ATCs have a survival rate of less than 1 year. More than half of patients have a previous or coexisting differentiated thyroid carcinoma. ATC presents with widespread local invasion and distant metastases in other organs of the body (Lewinski et al. 2000).

## **1.3 Molecular events involved with thyroid cancer**

There have been significant advances in the molecular mechanisms involved with thyroid neoplasia. The knowledge of genetic alterations which are implicated in the tumourigenesis provides the basis of the signalling activities within the thyroid tumour types. Of which, the MAPK pathway contains within it key signalling molecules which are initiating events in FA and PTC. The pathway and the oncogenic events identified in thyroid tumours are detailed below.

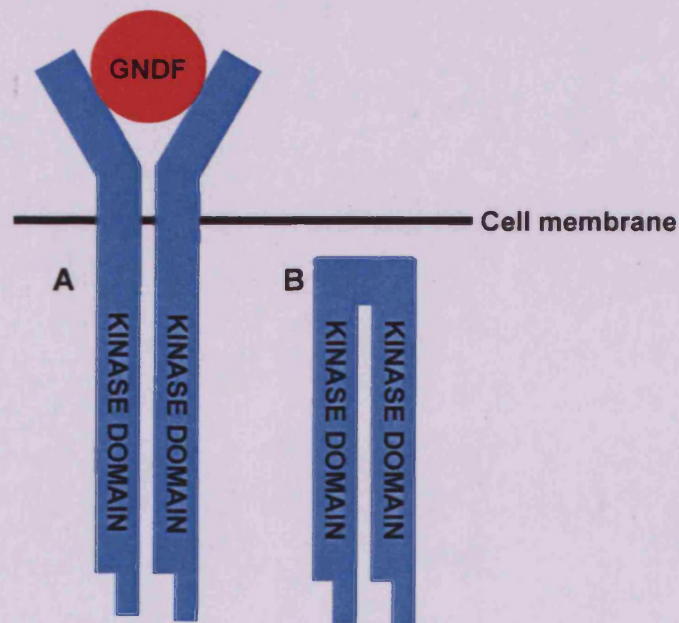
### **1.3.1 Mitogen-activated protein kinase (MAPK) signalling pathway**

The MAPK signalling cascade is a key pathway which leads to multifaceted cellular responses controlling cell proliferation, differentiation and survival in response to environmental conditions. The pathway is activated by extracellular stimuli through receptor tyrosine kinases to illicit a cellular response. Receptor tyrosine kinase activation triggers inactive Ras to exchange GDP to GTP to become active, and to recruit Raf to the plasma membrane. By phosphorylation, activation of MEK1 and MEK2 occurs which leads to the activation of ERK. This proceeds to the activation of numerous targets, which include kinases, phosphatases, transcription factors and cytoskeletal proteins. This pathway is particularly prevalent in cancer, with constitutive activation of the pathway caused by genetic alterations affecting the signalling cascade.

### 1.3.1.1 RET

RET (Rearranged during transfection) proto-oncogene is a transmembrane receptor tyrosine kinase localised on chromosome 10q11.2 and contains 21 exons. Its natural ligand is the glial cell derived neurotrophic factor (GDNF), and it normally functions in the development of the kidney and the enteric nervous system (Takahashi 1988; Schuchardt et al. 1994). RET activates multiple cell signalling cascades mediating a number of cellular functions including cell survival, differentiation, proliferation, migration and chemotaxis. RET chromosomal rearrangements are implicated in the oncogenesis of papillary carcinomas. Targeted expression of RET/PTC1 in transgenic mice developed PTC (Cho et al. 1999). RET expression is normally restricted to calcitonin-producing C-cells but not follicular cells of the thyroid. The fusion of intracellular RET tyrosine kinase domain to the N-termini of certain thyroid specific genes generates chimeric oncogenes leading to expression in follicular cells. These fusion proteins are designated RET/PTC. There are 12 RET/PTC rearrangements currently identified in papillary carcinomas (Ichiara et al. 2004) (Table 1.0). The most prevalent, RET/PTC1, involves the chromosomal inversion of the tyrosine-kinase domain of RET on chromosome 10q11.2 which juxtaposes with the promoter region and first exon of the gene H4(D10S170). The cloning and sequencing revealed that H4 shares no homology to any known genes. However it does contain a SH3 binding domain highlighting a role in protein interaction (Grieco et al. 1994; Celetti et al. 2004). Functional characterisation of H4 has highlighted its role in PTCs. It was found to have prominent roles in the formation of the RET heterodimer which is crucial for the constitutive activation and transforming ability (Tong et al. 1997). Restored expression of H4 in a RET/PTC1 positive cell line induced apoptosis. This indicates a role of H4 in cell survival and the pathogenesis

of cancer. RET/PTC rearrangements are localised in the cytoplasm where ligand independent dimerisation and constitutive activation takes place (Figure 2). RET/PTC rearrangements are found in 5-40% of PTC, and are more prevalent in sporadic PTC (13-44%) and cancers associated with exposure to radiation (50-90%) (Jhiang 2000).



**Figure 2.** Normal RET and oncogenic RET/PTC1. **A**, Normal RET. Upon GNDF ligand binding RET receptors dimerise and initiate signalling. **B**, Oncogenic RET/PTC1 fusion causes ligand independent constitutive phosphorylation of residues in the kinase domain. (Ichihara et al. 2004).

<b>Rearrangement</b>	<b>RET fusion partner</b>	<b>Mechanism and location</b>	<b>Reference</b>
RET/PTC1	H4 (D10S170 locus)	Inversion (chromosome 10) (q11.2;q21)	(Celetti et al. 2004)
RET/PTC2	PRKAIIRA	Translocation (chromosome 10:17) (q11.2;q23)	(Sozzi et al. 1994)
RET/PTC3 and RET/PTC4	Nuclear Receptor Coactivator 4 (NCOA4)	Inversion (chromosome 10) (q11.2)	(Bongarzzone et al. 1994; Santoro et al. 1994)
RET/PTC5	Golgi Autoantigen 5 (RFG5)	Translocation (chromosome 10:14) (q11.2;q)	(Klugbauer et al. 1998)
RET/PTC6	Tripartite Motif-containing 24 (TRIM24)	Translocation (chromosome 7:10) (q32;q11.2)	(Klugbauer et al. 1996)
RET/PTC7	Tripartite Motif-containing 33 (TRIM33)	Translocation (chromosome 1:10) (p13;q11.2)	(Klugbauer et al. 1996)
RET/PTC8	Kinectin 1 (KTN1)	Translocation (chromosome 10:14) (q11.2;q22.1)	(Salassidis et al. 2000)
RET/PTC9	RFG9	Translocation (chromosome 10:18) (q11.2;q21-22)	(Klugbauer et al. 2000)
ELKS-RET	ELKS/RAB6-interacting (ELKS)	Translocation (chromosome 10:12) (q11.2;p13)	(Nakata et al. 1999)
PCM1-RET	Pericentriolar Material 1 (PCM1)	Translocation (chromosome 8:10) (p21-22;q11.2)	(Corvi et al. 2000)
RFP-RET	Tripartite Motif-containing 27 (TRIM27)	Translocation (chromosome 6:10) (p21;q11.2)	(Saenko et al. 2003)

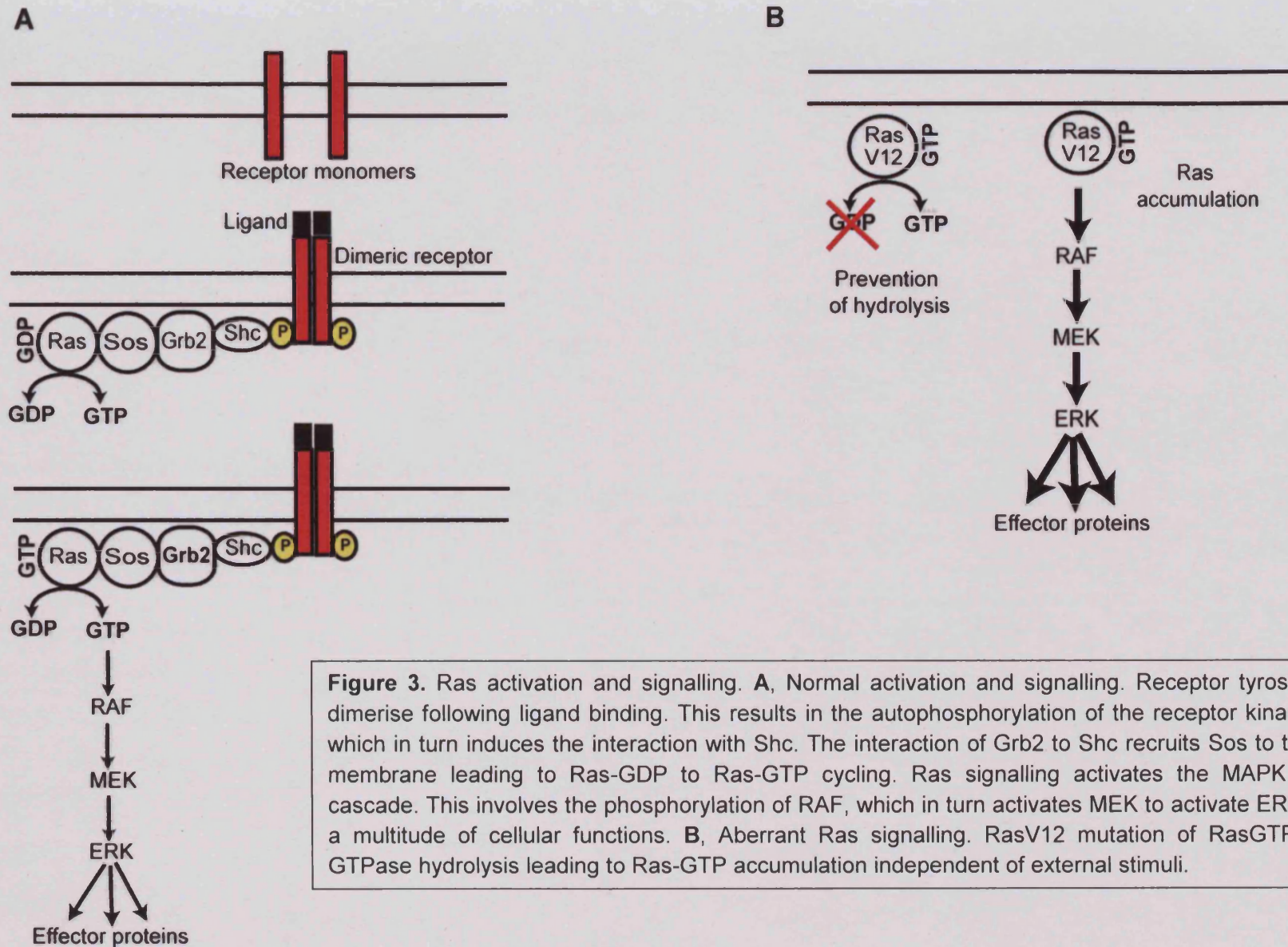
**Table 1.0.** RET rearrangements with fusion partners which are found in thyroid cancers. RET/PTC1 represents the most common rearrangement found in PTC (Putzer and Drost 2004; Arighi et al. 2005; Prensner and Chinnaiyan 2009).

#### 1.3.1.2 Ras

In humans, the RAS superfamily is categorised into 6 families, Ras, Rho, Arf, Rab, Ran and Rad. The Ras family are encoded by 3 Ras genes which comprise Harvey-Ras (H-Ras) located on chromosome 11, Neuroblastoma-Ras (N-Ras) located on chromosome 1 and two splice variants of Kirsten-Ras (K-Ras) which are located on chromosome 12. Their mechanism of activation involves cycling between an inactive state when bound to GDP, and an active state when bound to GTP (Takai et al. 2001). The GTP-GDP cycle is regulated by cell surface receptors comprising of receptor tyrosine kinases, heterotrimeric G protein-coupled receptors, cytokine receptors and integrins. Peptides like the Epidermal Growth Factor (EGF) binds to cell surface receptors resulting in receptor tyrosine dimerisation and activation of their kinase domains. This leads to their autophosphorylation and release of cryptic binding sites for the binding of proteins containing Src homology 2 (SH2) domains. The SH2 containing Shc protein is tyrosine-phosphorylated inducing its interaction with Grb2. This complex binds to the Ras nucleotide exchange factor Sos, forming the Shc-Grb2-Sos complex which localises to the plasma membrane. Mediated by the interaction of Shc with the phosphorylated receptor, cycling of GDP to GTP on Ras takes place (Figure 3A). The Guanine-nucleotide exchange factors (GEFs) promote the release of bound GDP to aid in the binding of GTP (Cherfils and Chardin 1999), whilst GTPase activating proteins (GAPs) deactivates the activity of Ras supporting the binding of GDP (Donovan et al. 2002). The active GTP state of Ras interacts with wide range of effector proteins, of which, the best characterised are the Raf family of kinases (Dhillon et al. 2002) and the phosphatidylinositol 3-kinases (PI3K) (Wymann et al. 2003). All Ras proteins undergo post-translational modifications which comprise of farnesylation of a cysteine residue in the C-terminus

of Ras, proteolysis of the last 3 amino acid residues and methylation of C-terminal carboxylic acid group (Giehl 2005). The H-Ras and N-Ras are palmitoylated at their final cysteine residues. These modifications allow Ras to localise to the inner side of the plasma membrane where they are active (Hancock 2003). Mutations at codons 12, 13 and 61 of Ras are found in 30% of all tumours. In thyroid lesions, Ras mutations are prevalent in ~85% in follicular adenomas (Vasko et al. 2003). Mutation of valine for glycine in codon 12 (designated RasV12) occurs within the GTP binding domain of the H-Ras isoform. The lack of dissociation of GTP causes accumulation of Ras-GTP, causing constitutive activity independent of external stimuli (Vasko et al. 2003) (Figure 3B).





**Figure 3.** Ras activation and signalling. **A**, Normal activation and signalling. Receptor tyrosine kinase dimerise following ligand binding. This results in the autophosphorylation of the receptor kinase domain which in turn induces the interaction with Shc. The interaction of Grb2 to Shc recruits Sos to the plasma membrane leading to Ras-GDP to Ras-GTP cycling. Ras signalling activates the MAPK signalling cascade. This involves the phosphorylation of RAF, which in turn activates MEK to activate ERK inducing a multitude of cellular functions. **B**, Aberrant Ras signalling. RasV12 mutation of RasGTP prevents GTPase hydrolysis leading to Ras-GTP accumulation independent of external stimuli.

### 1.3.1.3 RAF family

The first identified Raf oncogene, v-Raf functions as the transforming gene of the mouse sarcoma virus 3611 which induces fibrosarcomas and erythroleukemia in newborn mice and C-Raf (Raf-1) was identified as its proto-oncogene homolog (Rapp et al. 1983). A-Raf was next discovered from screening a mouse spleen cDNA library with a v-Raf probe. B-Raf, the final discovery is the homolog of the chicken C-Rml transforming gene in the retrovirus which causes spontaneous ovarian tumours (Eychene et al. 1992; Calogeraki et al. 1993)

In mammals, the Raf family of genes consist of A-Raf located on chromosome Xp11, B-Raf located on chromosome 7q32 and *Raf-1* located on chromosome 3p25. Each of these genes encode cytosolic proteins termed A-Raf, B-Raf and Raf-1 respectively. The Raf proteins share 3 conserved regions, termed CR1, 2 and 3. CR1 contains the Ras binding domain, which localises the kinase to the cell membrane, and a regulatory cysteine rich domain. The CR2 domain is less defined but is rich in serine/threonine residues which aid in protein interactions. CR3 comprises the kinase domain and contains the activation loop and the negatively charged regulatory region (Figure 5). The Raf kinases differ based on their expression in tissues and kinase activity.

Mutational status of the RAF family of proteins in colorectal cancer revealed the B-Raf gene was frequently mutated whilst analysis of A-Raf and Raf-1 revealed rare germline polymorphisms, the majority of which were intronic and therefore unlikely to be important in the development of colorectal cancers (Fransen et al. 2004). In addition, mutational status of A-Raf in human cancers including ovarian epithelial tumours, colorectal carcinomas, gastric adenocarcinomas and acute leukemias failed to identify any A-raf mutations in the tumour samples (Lee et al. 2005). These

findings allude to the observation that B-Raf is the key ERK activator in the MAPK pathway (Mercer and Pritchard 2003). Indeed, in a resting state, Raf-1 is localised in the cytoplasm where it is phosphorylated on S259 and S621 inducing binding to 14-3-3 proteins to retain an inactive Raf-1 conformation. For Raf-1 activation, the phosphorylation of Y341 and S338 and de-phosphorylation of S621 are required (Mercer and Pritchard 2003). The constitutive activation of S446 in B-Raf (constituting to S338 in Raf-1) together with the D448 contributes to the high level of basal kinase activity which is 50-20 times higher than Raf-1 (Mason et al. 1999). This also means the activation required for A-Raf and Raf-1 is much more complex than BRAF. This also highlights why B-Raf is selected as an oncogene whilst A-Raf and Raf-1 are not.

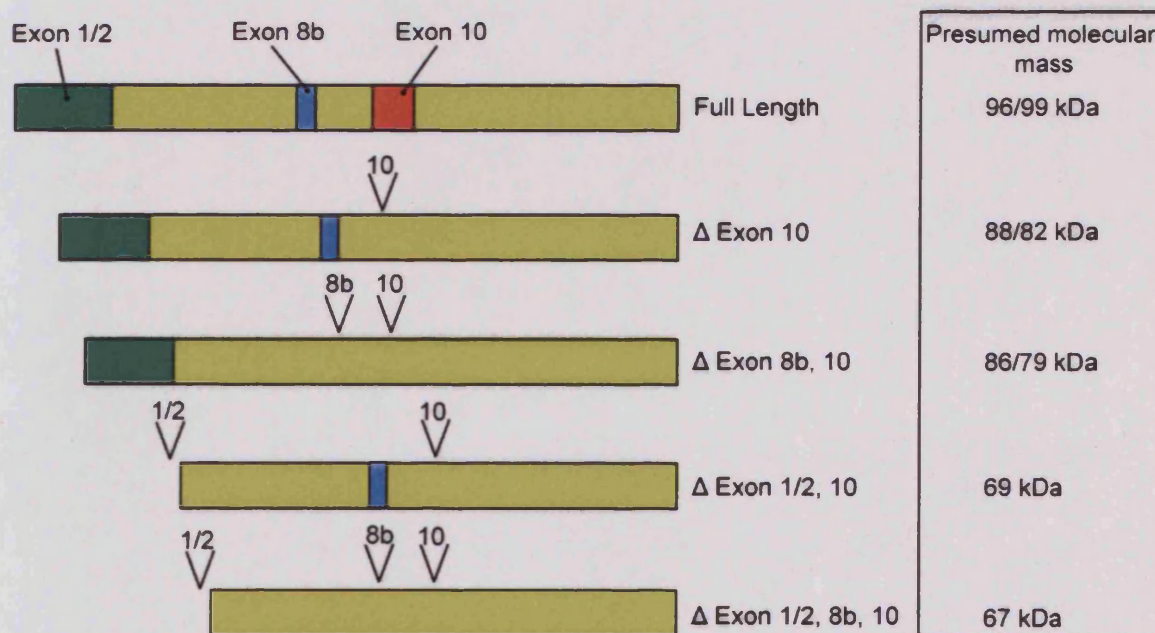
Expression of B-Raf is restricted to neural, testicular, splenic, and hematopoietic tissues. Activation of B-Raf is dependent upon Ras-GTP binding and 14-3-3 displacement. The displacement of 14-3-3 is induced by phosphorylation of S259 leading to an open conformation and exposure of the MEK binding site (Dhillon et al. 2002). In comparison, Raf-1 and A-Raf undergo a complex series of phosphorylation events to become active, whereas B-Raf has a constitutively active S445 residue. Binding of Ras-GTP at the cell membrane phosphorylates B-Raf inducing an active conformation. The phosphorylation of S365, located in the CR2 domain and S429 induces a downregulation of B-Raf activity by inducing the docking of 14-3-3 proteins (Guan et al. 2000).

Alternative splicing of B-raf gives rise to multiple isoforms (Figure 4). In mice, alternative splicing encodes at least 10 isoforms of B-Raf with tissue-specific expression. The isoforms arise from the alternative splicing of exon 10 (9b), exon 8b and exons 1/2. The presence of exon 10 increases the basal kinase activity and



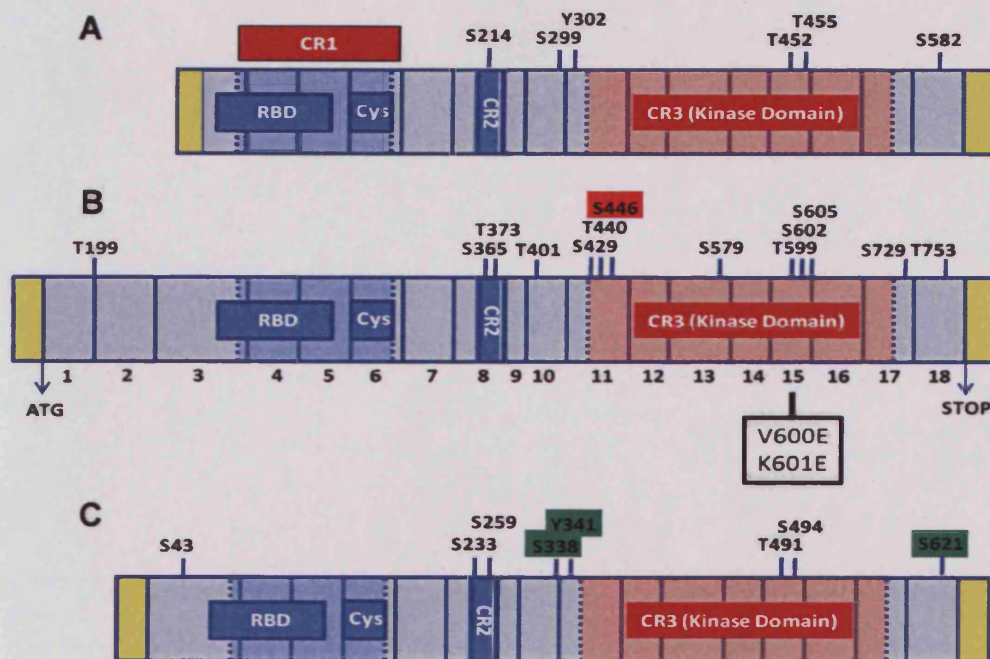
affinity towards MEK, whilst exon 8b decreases kinase activity (Hmitou et al. 2007). Exons 1/2 promote dimerisation with Raf-1 providing an additional level of kinase activity. These spliced exons modulate the kinase activity of B-Raf which may relate to its biological activity.

Signalling of the MAPK pathway through B-Raf induces sustained ERK activation and cellular differentiation by phosphorylation of dual-specificity kinases MEK1 and 2 which phosphorylate ERK1 and 2. The phosphorylation events shuttles ERK to the nucleus which induces the multitude of biological effects. Occurring in the cytosol, an additional level of activity is provided by the heterodimerisation of B-Raf with Raf-1, in a Ras dependent manner (Weber et al. 2001; Garnett et al. 2005). In this instance the activation of Raf-1 requires activation region and S621 phosphorylation (Figure 5).



**Figure 4.** Schematic representation of B-Raf isoforms. Spliced exons are highlighted by black triangles. Exon 1/2 is highlighted in green, exon 8b in blue and exon 10 in orange. Apparent molecular mass of the B-Raf isoforms was previously determined by a different group. This was determined by immunological characterisation with the different sera from experiments using transfected COS-1 cells or mouse tissues (Barnier et al. 1995).





**Figure 5.** Schematic diagram of the RAF family of proteins, displaying conserved domains with phosphorylation sites. Green boxes signify activation steps. Red box is constitutively active. Mutations of B-Raf are signified by a Black box. **A.** A-Raf **B.** B-Raf and **C.** Raf-1.

B-raf remains the only Raf kinase to be frequently mutated in human cancer with the highest incidence occurring in malignant melanoma (27%-70%), papillary thyroid carcinoma (36%-53%) and colorectal carcinoma (5%-22%) (Garnett and Marais 2004). Low frequency mutations are also present in a variety of other cancers (1%-3%) (Garnett and Marais 2004). Missense mutations of B-Raf have been identified in 24 different codons, and a majority of these are in low occurrence. Only 2 mutations have been identified in human thyroid lesions, the K601E and V600E (Figure 5). The K601E B-Raf mutation is rare and was only found in 3 cases follicular-variants of papillary carcinoma (Trovisco et al. 2004) and 2 cases of benign adenoma (Soares et al. 2003). Situated between T599 and S602, the K601E mutation has 2.5x lower kinase activity than the V600E mutation. The more prevalent mutation is the T1799A transversion which induces the V600E mutation of B-Raf. This mutation accounts for

~90% of all B-Raf mutations in thyroid cancers (Garnett and Marais 2004). This mutation mimics T599/S602 phosphorylation rendering B-Raf constitutively active. Transfection of BRAF<sup>V600E</sup> in NIH3T3 cells transforms cells approximately 667-fold more efficiently than wild-type B-Raf and induces constitutive ERK activation (Davies et al. 2002). In PCCL3 rat thyroid cells, BRAF<sup>V600E</sup> induces DNA synthesis, apoptosis, dedifferentiation and chromosomal instability (Mitsutake et al. 2005). Maintenance of proliferation and transformation by BRAF<sup>V600E</sup> was demonstrated in human papillary carcinoma cell lines (Liu et al. 2007). The BRAF<sup>V600E</sup> mutation is associated with aggressive characteristics including lymph node metastasis, extrathyroidal invasion and recurrence (Namba et al. 2003; Palona et al. 2006). Interestingly, the BRAF<sup>V600E</sup> mutation is also present in occult micro-papillary carcinomas (Ugolini et al. 2007). These lesions are designated <10mm in size and are typically discovered post-operative procedures for benign lesions following histological examination. Microcarcinomas exhibit features consistent with papillary carcinomas, but little is known about their potential to progress into papillary carcinoma. This indicates a prominent role of the B-raf oncogene, and in particular the BRAF<sup>V600E</sup> mutation in the initiation and progression of papillary carcinoma.

## **1.4 Gene expression profiling of thyroid lesions**

Microarrays are rapidly advancing, powerful tools to simultaneously analyse gene expression of thousands of genes. Although an expensive technology, it remains a widely utilised tool in research. There are many perspectives, considerations and limitations for microarrays, (reviewed by Eszlinger et al. 2007) and it has been used to elucidate various aspects of thyroid tumourigenesis. This chapter will review recent studies which use the microarray platform to gain new perspectives on malignant and benign thyroid lesions, highlighting aims, methods, results, perspectives and limitations.

### **1.4.1 Differentiating Thyroid Tumours**

The best method of diagnosis remains the fine needle aspiration biopsy which can diagnose papillary carcinoma with an accuracy of 95%. However, tumours of follicular origin (FA and FTC) are not as easy to distinguish. Suspicious lesions are subjected to thyroidectomy, and subsequently histologically examined for capsular invasion. This forms the basis of many thyroid studies which use the microarray platform to distinguish FA from FTC. The microarray expression profiling offers a means to detect markers to distinguish tumour types together with key biological background information.

Global gene expression analysis offered a means of dissecting the minimal number of genes that have a role in FTC progression, providing new markers for therapy and differentiating FTC from FA (Weber et al. 2005). With sample consisting of 12 FTCs and 12 FAs, this study identified 96 probe sets corresponding to 80 genes, which

were categorised based on their histological groups. Following fold-change ranking of the probe sets, 11 genes were returned as significant. The top two ranking genes confirmed by quantitative RT-PCR were Cyclin D2 and growth differentiation factor 15. Interestingly, these genes have not previously been associated with thyroid carcinogenesis and thus remain unique to this study. This study showed like other studies that FTCs can be differentiated by a select number of genes (Aldred et al. 2004). The Aldred et al. study differentiated FTC from PTC using 5 genes. The structure of the experiment was conducted by profiling FTC and PTC independently and combining the datasets to highlight the similarities and differences. Compared to normal thyroid, PTCs displayed two distinct clusters of genes including Cbp/p300-interacting transactivator 1, Claudin-10 and Insulin-like growth factor binding protein-6. FTCs did not express Claudin-10 and had a decreased expression of Insulin-like growth factor binding protein-6, Caveolin-1 and Caveolin-2. The datasets were analysed using hierarchical clustering and analysed using *t*-tests to observe the overlap in change calls and gene expression levels of PTC when compared with FTC. Consistent with other studies, galanin3 (GAL3) was over-expressed in both PTC and FTCs (Lubitz et al. 2006). Although these genes are differentially expressed in the tumour types, their potential roles as prognostic and therapeutic markers remain to be defined.

An early study used a cDNA microarray platform containing 5,760 probes for diagnostic and clinical implications profiling 6 normal thyroid samples, 4 FAs, 3 FTCs and 2 PTCs, and 3 follicular variant of PTC (FVPTC) samples (Chevallard et al. 2004). This study showed consistency with a number of other studies. There were 165 differentially expressed genes in FA, 222 in FTC and 172 in PTC. The biological function of the 165 genes were classified into categories of cytoskeleton, cell



adhesion, cellular matrix related, signal transduction and immune-response, a finding consistent with other profiling studies (Giordano et al. 2005; Melillo et al. 2005; Delys et al. 2007; Vasko et al. 2007). The data also showed that FA can be differentiated for FTC by a set of 43 genes, 10 of which were previously reported (Barden et al. 2003). A set of 23 and 52 genes could differentiate PTC from FA and FTCs. These genes were composed of members belonging to protein metabolism/catabolism (15%), intracellular trafficking (12%), ion/metal binding and transport (10%) and transcription (12%). In general, the expression signatures of PTC and FTC showed common features, but individual genes including dual specificity phosphatase 6 (DUSP6), dual specificity phosphatase 5 (DUSP5), milk fat globule-EGF factor 8 protein were identified as differentiation markers between the tumour types, and this is a finding repeated in other studies (Delys et al. 2007). Many of these genes have not been associated with thyroid development and may lend themselves to potential diagnostic and prognostic markers. In summary, this study showed gene expression data correlated with classification of thyroid lesions, and identified genes that can differentiate between the tumour types. From a technological stand point, the microarray platform used contained very few probes compared to today's platforms. Microarray platforms are able to differentiate tumour types, however differences in platform and analysis makes each study unique, and it is difficult to directly compare similar studies.

Improvement on the methods for diagnosing malignancy from fine needle aspirates is an issue which required further investigation. Instead of tumour tissue samples, which require invasive procedures FNA biopsies were profiled on a microarray platform with an aim to assess the feasibility for diagnosing malignancy from FNA (Lubitz et al. 2006). To elucidate this aim, tissue samples were profiled and

compared with FNA. The tumour samples included FA and hyperplastic nodules which were compared with PTC and FVPTC. Profiling was conducted on the Affymetrix U133 gene chips using tissue samples. Differential expression revealed a set of 25 genes which included thyroid peroxidase, transcription termination factor RNA polymerase III, GAL3, fibronectin (FN), TIMP metalloproteinase inhibitor 1 (TIMP1) and keratin 19, which correlated with other studies (Huang et al. 2001; Giordano et al. 2005; Prasad et al. 2005). The FNA categorised as either preoperative diagnosis or as a final histological diagnosis were compared. Of these cases, 16 of the 17 were predicted by the microarray to be benign or malignant which was confirmed by pathological analysis. Preoperative analysis was less conclusive with 13 of 17 (76%) of cases confirmed to be correctly diagnosed. This study represents a breakthrough for using microarray technology as a platform to improve diagnosis. More importantly, the study shows consistency of the Affymetrix microarray platform, the use of alternative sample material, correlation with other studies and the use of a microarray as a useful diagnostic tool albeit its high cost and technical complexity.

There are atypical thyroid tumours which are difficult to determine by histological analysis, in particular the diagnosis of FVPTC which represents a challenge as it retains features of both benign and malignant lesions. An expression profile was generated for malignancy in PTC to identify discriminating genes to distinguish the subtypes of PTC (Finn et al. 2007). A differentially cluster of 173 genes were identified from the datasets. In conjunction with other studies genes including soluble galactoside-binding lectin 3, v-yes-1 Yamaguchi sarcoma viral related oncogene homolog, TTF3, cellular retinoic acid binding protein 1, BCL2-associated X protein, mitogen-activated protein kinase 4, CD44, TIMP1, fibroblast growth factor receptor 2

and S100A11 also displayed significant expression (Delys et al. 2007). The unsupervised clustering of FVPTC and PTC did not reveal any independent clustering highlighting the close relationship between the tumour types. However, 15 genes were differentially expressed in FVPTC compared to benign and classical PTC. The cathepsin family of genes were all highly upregulated in FVPTC. The authors highlight the close biology of PTC and FVPTC, however the identification of a cohort of genes in FVPTC may improve diagnosis of this lesion. This study has similarities with other unique studies where proteases including MMP-2 have been highlighted.

#### **1.4.2 Molecular characterisation of thyroid tumours**

Detailed molecular characterisation of thyroid tumours have been a primary aim for a number of studies. One such study was performed to determine novel clinical markers (Zhao et al. 2006). Gene expression profiles using filter-based high-density cDNA arrays in a series of thyroid tissues which included 17 normal thyroid, 16 PTCs, 13 FTCs. These cDNA arrays contained 12,000 known genes and 63,000 expressed sequence tags (EST). ANOVA analysis revealed 1,950 cDNA targets displaying significant expression differences between FTC, PTC and normal thyroid. Occurring in >40% PTC and/or FTC, 603 genes were 2-fold under-expressed and 176 genes 2-fold over-expressed. Despite the differential expression of the tumours when compared to normal thyroid, there were minimal differences seen between FTC and PTC. The most frequent over-expressed genes in >60% of FTC and/or PTC were functionally categorised into cell cycle control, signal transduction, growth factor receptors (EGF receptor) transcriptional factors and cell interaction. These

processes are consistent with involvement of protease activity identified in other studies (Melillo et al. 2005; Delys et al. 2007; Vasko et al. 2007). Consistent with other studies thyroid specific markers including thyroglobulin and paired box 8 were under-expressed in both PTC and FTC (Melillo et al. 2005; Delys et al. 2007). FN had increased expression in 81% of PTCs and 33% of FTCs and BSG was over-expressed in FTC but not in PTC (Prasad et al. 2005; Vasko et al. 2007). Significant expression variation revealed 123 genes (45 known, 78 unknown) which were identified to be differentially expressed between PTC and FTC. The most differentially expressed genes identified by permutation 2-sample *t*-test were over-expressed in PTC and under-expressed in FTC. They concluded the similar gene expression profile between PTC and FTC reflects a close histological and biological relationship. However, the differences between PTC and FTC may contribute to the progression of the carcinomas along different pathways.

PTCs are commonly associated with lymph node metastases, and this occurs in 30% to 65% of cases at initial diagnosis. PTCs with metastasis are phenotypically more aggressive, and are characterised by lymphatic invasion, increased resistance to treatment and increased mortality. Microarray profiling has been used to identify biomarkers related to PTC with aggressive metastases (Cerutti et al. 2007). This analysis also highlighted invasive processes and genes involved with PTC. Using an individual patient with metastatic PTC, serial analysis of gene expression (SAGE) was conducted due to its ability to produce consistent gene profiles from small samples, and allows for easy comparison with data from other studies. Retrospective comparison of the data obtained from 4 different SAGE libraries revealed 31 out of 498 transcripts were highly expressed in the metastatic libraries but not in the patient tumour, whilst 47 transcripts were highly expressed in the

tumour but not in the libraries. Transcripts were chosen based on their fold induction were chosen from the libraries, as they have the greatest potential as prognostic markers, together with S100 calcium binding protein A11 (S100A11), LSM7 and met proto-oncogene (cMET) which have been previously implicated in tumours with metastatic potential. These were validated using quantitative PCR and compared with the levels associated with normal thyroid physiology. The gene ontology was related to transport and cell signalling. The SAGE and quantitative PCR were compared using the matched samples of normal thyroid, primary PTC and lymph node metastasis to generate libraries. Specifically, protein tyrosine phosphatase receptor type C was found to be over-expressed in all metastases analysed. LIM domain containing 2 (LIMD2) and the CD48 molecule were only expressed in lymph node metastases whilst lymphotoxin beta and ATP-binding cassette sub-family C member 3 were over-expressed in lymph node metastases, at lower levels in PTC and normal thyroid samples. The previously reported genes (S100A11 and cMet) were all found to concur with other studies (Finn et al. 2007). Although further investigation is required, the authors show that for the first time, LIMD2 is associated with the metastatic process of PTC. The authors also conclude the identification and validation of genes associated with lymph node metastases from PTC may be helpful in providing new targets for therapy, and earlier detection of metastasis. What is clear from this study to the identification of novel genes not highlighted in the other studies, which may be linked to the metastases and not the primary tumour, and this is the case for other tumour variants.

In parallel, PTC remains the most frequent endocrine malignancy and represents up to 80% of all malignant thyroid tumours (Delys et al. 2007). Gene expression profiling have provided tumour type signatures and elucidated tumour cell metabolism for

PTC (Baris et al. 2005; Giordano et al. 2005). The majority of gene profiling studies fail to explore the underlying mechanisms in detail. Microarray datasets from different experiments were combined to analyse the regulation of the numerous signalling pathways and processes involved in PTC (Delys et al. 2007). Their analysis generated a gene list with 451 upregulated and 233 downregulated genes. Gene ontology analysis revealed the genes to be related to immune response, cell adhesion, extracellular matrix, cell-cell adhesion, integrin, collagen, peptidase activity and the EGFR signalling pathway.

Hurthle cell carcinomas are malignant thyroid tumours which present with accumulations of dilated mitochondria. A hypothesis that high mitochondrial content may be linked to the malignant progression of PTC was investigated (Baris et al. 2005). The comparison between 6 oncocyctic follicular carcinomas and 6 mitochondrial-rich papillary carcinomas revealed a set of genes which can distinguish oncocyctic FTC from PTCs. Genes over-expressed in PTCs included DUSP6, cyclin D1, TIMP1, CD44 molecule (CD44) and matrix metalloproteinase 2 (MMP-2). In oncocyctic tumours, there were a number of different over-expressed genes. Gene ontology of the 59 PTC over-expressed genes was related to cellular physiological process (52.94%) and cell communication (43.14%). A total of 24 genes over-expressed in the oncocyctic follicular carcinomas were involved in the metabolism (83.33%) and stress response (21.05%). This suggested a higher rate of oxidative metabolism in oncocyctic tumours compared with PTCs. The respiratory chain complexes III and IV were over-expressed in oncocyctic tumours which may modulate the production of reactive oxygen species. They conclude that oxidative metabolism and reactive oxygen species distinguishes oncocyctic tumours from

PTCs. The over-expression of the respiratory chain proteins seems to be essential to the development of oncocytic FTCs.

The malignant phenotype of PTC is primarily caused by Epithelial to mesenchymal transition (EMT), and genes which regulate the process of EMT were profiled by analysing the invasive front of PTC (Vasko et al. 2007). They characterised the biological processes involved with PTC invasion and observed the expression of genes in invasive regions compared to central regions of papillary carcinomas. In addition to identifying BRAF<sup>V600E</sup> and RET/PTC genetic alterations, the invasive regions of PTCs revealed over-expression of genes relating to TGF $\beta$  cell signalling, integrin cell signalling, Notch signalling, NF $\kappa$ B signalling and PI3K signalling. Quantitative RT-PCR was conducted on a select number of genes known to be involved with EMT (FN, runt-related transcription factor 2 (RUNX2), prickle homolog 1 and slit homolog 1). Immunohistochemistry was conducted on RUNX2, FN and vimentin to assess the expression of the genes in PTC. Osteopontin (OPN) is also in part regulated by RUNX2, and expression of osteopontin was also confirmed in the PTCs. There are many invasive processes and genes which have been highlighted by these studies. Consistent with other studies they found the over-expression of genes including CITED1, FN (Delys et al. 2007), extracellular matrix protein 1 (ECM1), mucin 1 (MUC1) and S100 calcium binding protein A4 (S100A4) (Zou et al. 2005) in PTC, with reduced expression of genes regulating thyroid hormone production. Consistent with immunohistochemical studies (Shin et al. 2004), genes encoding activators of the JNK pathway were over-expressed in the dataset, suggesting a possible role of the JNK pathway. The over-expression of dual specificity phosphatase 4 (DUSP4), DUSP5 and DUSP6 was seen in the dataset (Chevallard et al. 2004). These are known inhibitors of the MAPK pathway may

account for the JNK pathway activation. Over-expression of many proteases was also observed. Known mediators which regulate the remodelling the extracellular included matrix metalloproteinase 1, matrix metalloproteinase 7 and matrix metalloproteinase 11, serine proteases cysteine proteases including urokinase plasminogen activator and aspartyl proteases including cathepsin B.

A number of studies have highlighted the genetic alterations of PTC in a microarray platform. The oncogenic Ras and B-Raf mutations with RET/PTC rearrangements are known to be mutually exclusive in thyroid lesions. The genetic alterations were genotyped in thyroid tumours (Giordano et al. 2005) and the link verified between the 3 oncogenes in a thyroid cell culture model (Melillo et al. 2005). This model uses a differentiated thyroid follicular cell line (PCCL3) derived from Fischer rats. A mixture of 6 hormones are required to induce proliferation and transfection of RET/PTC1, BRAF<sup>V600E</sup> and RASV12 to induce hormone independent growth of the cell line. Genes modified by cell transformation included the upregulation of G1 cyclins and downregulation of thyroid markers. 786 probes in total were upregulated in RET/PTC and 146 of these were found to be significant. Of these 146 probes 80 were also induced by RASV12 and 70 were induced by BRAF<sup>V600E</sup>. Overall 60 probes were induced by all three oncogenes and were related to structure and adhesion (S100A4 and S100 calcium binding protein A6) and proteolysis (matrix metalloproteinase 13, matrix metalloproteinase 10, matrix metalloproteinase 12 and matrix metalloproteinase 3). Additionally, the intermediate filament protein, vimentin was also upregulated, data which showed consistency with other studies (Vasko et al. 2007). Of 2517 probes, 338 probes were downregulated in RET/PTC1, 148 probes were downregulated in RASV12 and 121 probes were downregulated in BRAF<sup>V600E</sup>. Genes relating to thyroid differentiation (Thyroid peroxidase (TPO) and



thyroglobulin (TG)), adhesion, signalling and metabolism were also downregulated. To observe the role of the MAPK pathway in the expression regulation induced by RET/PTC3 transformed cells, quantitative PCR of 37 selected genes from the genetic profiling data revealed that with the use of the MEK inhibitor U0126 or siRNA directed towards BRAF, 22 of the 37 genes were altered by BRAF knock-down. This highlighted the integral role of the MAPK pathway in the expression profile of transformed cells. In addition, there are a number of genes which overlap between cells transformed by RET/PTC3, BRAF<sup>V600E</sup> or RASV12, and most of these genes were regulated by the MAPK pathway.

As a comparison, combined transcriptional expression profiles of 51 papillary carcinomas with their known morphology and their RET/PTC, B-Raf and Ras mutational status was analysed in order to improve the understanding of papillary carcinoma and the relationship between the gene expression profiles, morphology and oncogenic events (Giordano et al. 2005). The relationship between mutation and gene expression revealed distinct gene clusters which were dependent on RAS or BRAF mutation or RET/PTC rearrangement. There was a strong correlation between mutation and gene expression compared to the gene expression and morphology. Application of increased stringent analysis criteria resulted in a total of 132 (100 upregulated, 32 downregulated) probe-sets of RET/PTC, 82 (31 upregulated, 51 downregulated) probe-sets for B-Raf and 571 (165 upregulated, 406 downregulated) probe-sets for RAS. This indicated the increased number of downregulated genes for RAS with a similar number in B-Raf and RET/PTC indicates the thyroid carcinoma pathologies are dependent on the oncogenic event. The mutation specific expression of cMet and TIMP1 was confirmed in PTC. Genes relating to immune response were also differentially expressed in a mutation specific manner. The expression of TPO

was decreased 18.9 fold in PTC compared with normal thyroid, a result consistent with another study (Melillo et al. 2005). The authors conclude the reduced TPO expression suggests decreased thyroid hormone synthesis and radioiodine uptake, increasing the resistance of the B-Raf mutant tumours to radioiodine treatment. The greatest difference was found in B-Raf tumours compared to Ras and RET/PTC tumours. This may relate to the specific signalling cascades involved with the oncogenic events as B-Raf mutations follow a more aggressive course. Since the aims of these particular studies are unique and utilise different sample material, the datasets cannot be direct compared but they do indicate a number of similarities in gene expression with thyroid differentiation, proteolysis and immune response all consistent throughout.

#### **1.4.3 Genetic Profiling of thyroid tumour cell lines**

Thyroid tumour cell lines derived from different pathologic origin have been developed over past years. The unlimited replicative potential and the existence of only one cell type allows for a more accurate interpretation. The *in vitro* properties of thyroid tumours were investigated by using thyroid tumour derived cell lines which included 1 FA (KAK-1), 2 FTC (FTC133 and WRO), 3 PTC (BCPAP, TPC-1 and KAT-10) and 2 ATC (KAT-1 and 8505C) (van Staveren et al. 2007). The gene expression of the cell lines were compared with normal human thyrocytes from primary cultures. The hierarchical clustering confirmed by bootstrap analysis, revealed the cell lines did not cluster based on their origin. The KAK-1, KAT-4 and KAT-10 cell lines were highly similar, all containing a point mutation in E-Cadherin and a heterozygous BRAFV600E mutation. Regardless of origin, comparison of

these cell lines with a panel of *in vivo* solid tumours showed the cell lines resembled *in vivo* ATCs. The gene expression profiles of the *in vivo* tumours showed that the adenomas clustered together but were distinct from ATCs. The FTCs and PTCs clustered in between adenoma and the ATCs. RT-PCR confirmed the mRNA expression of thyroid differentiation markers. The cell lines were found to have lost the expression of a number of thyroid markers (thyroid stimulating hormone receptor (TSHR), sodium iodide symporter member 5 (NIS), TG and TPO). The authors conclude the cell lines cultured in the absence of thyroid stimulating hormone may account for partial loss of differentiation. This also highlights the limitations of using cell lines in a gene profiling study as the study raises the issue of origin of the cell line. They also speculate that cell lines derived from *in vivo* tumours may have arisen from cancer stem cells or adapt to the *in vitro* environment which they are cultured in.

Other models discussed in this chapter include the transfectable rat cell line PCCL3 model (Melillo et al. 2005), which would circumvent the background noise with primary tissue profiling but being derived from a different species, raised the question if this would alter the gene expression. The use of cell lines derived from human tumour tissue would provide a species-relevant expression profile and being a pure culture would minimise the background noise seen with tumour tissue samples. However, their transformed nature and unlimited proliferative potential may affect the gene expression. In addition, evidence has emerged that the majority of cell lines resemble anaplastic tumours regardless of where they were derived from (van Staveren et al. 2007). Clearly, there are positive and negatives associated with all the different sample material for profiling, but raise the question of consideration of selection of the sample material to create an accurate and precise study reflecting

the true tumour situation. The majority of profiling studies discussed all have overlapping genes, pathways and processes which are known to be involved in thyroid tumourigenesis, and have been used to assess the validity of a given study. The involvement of individual genes would provide the greatest confirmation of precision and accuracy and this is an adopted method of confirmation by a number of studies.

#### **1.4.4 Overall perspectives**

A key goal of many studies is to distinguish between tumour types and to provide detailed molecular characterisation. This essentially provides an insight into tumour progression and identifies potentially markers to facilitate diagnosis and thereby guide treatment. Essentially, each microarray gene profiling study is unique based on the type of sample, the platform and the analysis method utilised. Despite these variables there is consistency between the studies. For instance a number of genes were identified as distinguishing markers. GAL3 was found overexpressed in both FTC and PTCs (Aldred et al. 2004; Lubitz et al. 2006). FN was also identified in a number of studies. This gene was overexpressed in a large percentage of PTCs and FTCs (Prasad et al. 2005; Vasko et al. 2007; Zhao et al. 2006). S100A4, a marker of thyroid malignancy was also identified in PTCs (Giordano et al. 2005; Melillo et al. 2005; Zou et al. 2005). A number of studies also identified TIMP1 malignant thyroid tumours (Chevallard et al. 2004; Vasko et al. 2007; Baris et al. 2005; Lubitz et al. 2006). The dual specificity phosphatase were also identified in a number of studies (Chevallard et al. 2004; Vasko et al. 2007).

There is a clear indication that despite differences in methodology there is some consistency between studies. However, the type of sample, the platform and analysis method ultimately has an impact on the profiling data, and this requires careful consideration for future studies.

## 1.5 In vitro models of thyroid neoplasia

My group have established *in vitro* models representing FA and PTC. Normal primary thyrocytes have a low proliferative rate (3-4 population doublings), and retroviral induction of mutated H-RASV12 enhances proliferation and develops a phenotype consistent with *in vivo* benign follicular adenomas. The enhanced proliferation is not definite in thyrocytes, with growth ceasing after 15-25 population doublings, resembling replicative senescence (Bond et al. 1994; Jones et al. 2000). This finding is consistent with *in vivo* data as FAs appear to reach a self-limiting quiescent state. Retroviral transduction of RASV12 results in a distinctive expanding pattern of growth with distinct colony margins, a phenotype consistent with FA. Conversely, the retroviral transduction of RET/PTC1 and BRAF<sup>V600E</sup> also results in a self-limiting growth response, with the resulting colonies showing a different pattern of growth. The RET/PTC1 and BRAF<sup>V600E</sup> cells are capable of migrating over and between the surrounding cells, a phenotype reminiscent of *in vivo* PTC.

## 1.6 Aims, Objectives and Experimental strategies

Thyroid tumours are diagnosed with pathological examination of FNA, an invasive procedure. This diagnostic technique is not absolute and is unable to differentiate invasive FTC from benign FA.

Much work has been conducted to identify additional markers of malignancy. Microarray gene profiling offers an alternative method for defining malignancy in thyroid tumours. Currently, microarray gene profiling has been used to characterise malignancy, differentiate tumour types and define biological information in thyroid. These studies either use cell lines which have restricted biological complexity or primary tumours which possess a high signal to noise ratio.

My group has established *in vitro* primary models of PTC and FA which represent the *in vivo* situation. The infection of the RASV12 oncogene in thyroid epithelial cells results in a phenotype consistent with FA. Conversely, introducing the BRAF<sup>V600E</sup> and RET/PTC1 oncogenes results in a phenotype consistent with PTC. These models represented a major incentive to study malignant behaviour in thyroid cancer using a microarray platform. Firstly, this is the only study to utilise this type of sample in a microarray platform. Secondly, the *in vitro* primary models offer a closer representation to *in vivo* tumours compared to cell lines and on a microarray platform, a lower signal to noise ratio compared to primary tumour samples.

The major aims of the thesis were to use a microarray platform using the previously established *in vitro* primary models of PTC and FA to identify novel markers associated with tumour invasion. I next aimed to investigate if the markers were of therapeutic value by genetic manipulation. These aims would not only elucidate the biological background of thyroid tumours in relation to tumour invasion and

progression, but also provide evidence of additional therapeutic markers which could be used in diagnostic and therapeutic approaches.

The following are the major aims of the thesis:

### **Identification of novel markers of tumour invasion using a microarray platform**

#### **1) Defining the $BRAF^{V600E}$ mutation in thyroid cell lines**

The B-Raf oncogene is known to be alternatively spliced giving rise to multiple isoforms. I wanted to define these isoforms in thyroid cancer cell lines and analyse the B-Raf oncogenic mutation in the isoforms. This was of importance to future investigations which use the  $BRAF^{V600E}$  oncogene. To achieve this aim the following objectives were accomplished:

1. Characterisation of B-Raf isoforms in thyroid cell lines
2. Define the mutational status in thyroid cell lines

#### **2) Identification of genetic markers involved in tumour invasion**

The microarray platform offers genome-wide analysis of gene expression. Profiling of the *in vitro* models not only accomplished the primary aim of identifying novel markers which define an invasive phenotype but also provided relevant biological information. To accomplish this aim the following objectives were carried out:

1. Infecting primary epithelial cells with the RASV12,  $BRAF^{V600E}$  and RET/PTC1 to formulate the primary *in vitro* models



2. Processing primary models for the use in a microarray platform
3. Conduct the genetic profiling using the microarray platform
4. Analysis of the datasets to identify a gene cluster implicated in invasive behaviour
5. Identify genetic markers using a literature analysis

## **2.0 Materials and Methods**

This material and methods section describes the assay procedures for producing the *in vitro* models and the one-cycle eukaryotic labelling of gene chips. After I extracted total RNA from samples the one-cycle eukaryotic labelling for the Microarray analysis was conducted by the in-house Microarray service provided by the Cardiff University CBS department.

## **2.1 Generation of oncogene expression infected thyroid colonies**

### **2.2.1 Cell culture**

Monolayer cultures of primary human epithelial cells were previously prepared by from surgical samples of histologically normal thyroid tissue by collagenase B and Dispase II (Roche, Germany) protease digestion with mechanical disaggregation (Williams et al. 1988). Cultures were judged as >99% epithelial by cytokeratin staining (Bond et al. 1994). Primary cultures were placed in liquid nitrogen storage until ready for use.

Frozen primary epithelial cells were thawed and maintained in a 1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 supplemented with 10% foetal calf serum (Autogen Bioclear, London, UK), Penicillin and Streptomycin (Sigma–Aldrich, UK) with glutamine. Cells were grown in a humidified atmosphere containing 3 % O<sub>2</sub>, 5 % CO<sub>2</sub> and 92% N<sub>2</sub> at 37°C.

### *2.2.2 Preparation of viral supernatants*

All procedures during viral infection were undertaken with significant caution with accordance with the advisory committee on Genetic Modification (ACGM) guidelines. All procedures involving the transfer and filtration of viral supernatants were conducted using a sterile plastic kwill and syringe in a class II containment facility. All pipette tips and stripettes were rinsed with bleach after use. All viral waste was placed in a special designated metal container and autoclaved before disposal.

#### *2.2.2.1 Preparation of Retroviral supernatants*

Ψ-CRIP-DOEJ producer cells were used to produce retrovirus encoding RASV12. Ψ-CRIP-RET/PTC1 producer cells were used to produce retrovirus encoding RET/PTC1. Both constructs have been previously described (Bond et al. 1994). For Ψ-CRIP-DOEJ and Ψ-CRIP-RET/PTC1 cells were grown to 95% confluency in selective medium (Table 2.0). 24 hours prior to infection the medium was removed and the cells replenished with a minimal amount of selective-free medium appropriate for the target cells containing freshly added glutamine. This medium containing the retroviral particles was harvested for retroviral infection by collecting with a pipette and kwill into a 15mL centrifuge tube and centrifuging at 100 rpm for 5 minutes to sediment any cell debris.

For retrovirus encoding BRAF<sup>V600E</sup> the phoenix producer cell line was used (Morgenstern and Land 1990; Kinsella and Nolan 1996). To prepare BRAF<sup>V600E</sup> retrovirus phoenix cells were plated at a density of  $8 \times 10^6$  per T<sub>75</sub>- flask (Nalge Nunc, Intl.). After 24 hours, 10μg of BRAF<sup>V600E</sup>pbabeneo plasmid DNA was diluted in a total

of 405µL of sterile water. Using the calcium chloride method of transfection the plasmid DNA was introduced into the phoenix cells.

This involved re-feeding phoenix cells with 15 mL of medium and in a separate microtube adding 45µL of CaCl<sub>2</sub> to the diluted DNA, mixing and add drop wise to the HEPES-buffered saline solution (HeBs) while passing bubbles through the (HeBs). The mixture was vortexed for 2 seconds and left to incubate at room temperature for exactly 20 minutes. Fifteen minutes into the incubation, chloroquine at a concentration of 25µM was added to the phoenix cells. The transfection mix was then added after the final 5 minute incubation. The mixture was added whilst swirling the flask and incubated overnight in an incubator set at 37°C overnight. The following day the flasks were incubated at 35°C overnight. The next day the supernatant containing the BRAF<sup>V600E</sup> retrovirus was harvested. The supernatant was carefully removed and spun at 1000 rpm for 5 minutes to pellet any cell debris. The supernatant was then aliquoted and flash frozen in liquid nitrogen and stored at -80°C.

### 2.2.3 Viral infections

Primary epithelial were plated at 5x10<sup>5</sup> per 60mm-dish without selection. After 2 days, 1 hour prior to infection cells were refed with medium containing 8µg/mL polybrene. The dishes were aspirated to remove the medium and replaced with the retroviral supernatant containing 8µg/mL polybrene. After 5 hours, additional medium was added to the dishes, where the dishes were returned to the incubator. After 24 hours, cells were refed with non-selective medium and returned to the incubator. After 3 days, cells were passaged using non-selective medium into fresh dishes. After 24 hours, the dishes were refed with selective medium. Every 2 days dishes

were refed with selective medium for 3 weeks at which point colonies had been established.

Oncogene/plasmid	Producer cells	Selection
Ψ-CRIP-DOEJ-RASV12	ΨcripDOEJ	G418
BRAF <sup>V600E</sup> pbabeneo	Phoenix	G418
Ψ-CRIP-RET/PTC1	ΨcripRET/PTC1	G418

**Table 2.0.** Retroviral plasmids, producer cells and selection

## 2.3 Total RNA extraction

Cells were grown to 70% confluency in 100mm dishes and washed with pre-warmed Hanks Balanced Salt Solution (Sigma-Aldrich, Dorset, UK). To each dish, 1.5mL of TRIzol reagent (Invitrogen, Paisley, UK) was added and left to incubate for a maximum of 30 minutes to lyse the cells. Cells were scraped from the 100mm dishes using a Nunc cell scraper (Nalge Nunc, Intl.). The cell lysates were pipetted and aspirated to ensure homogenisation. The fractions were divided into 175µL aliquots and transferred into 1.5mL RNase-free centrifuge tubes, to which 175µL of chloroform was added. In parallel, Heavy Phase Lock Gel™ (PLG; Eppendorf) tubes were equilibrated by centrifugation at 13,000g for 25 seconds. The TRIzol/chloroform mixture was transferred to the phase lock gel tube and centrifuged at 13,000g for 2 minutes. The upper aqueous solution was transferred to a fresh eppendorf tube and the RNA precipitated by addition of 500µL of isopropanol and 5mg/mL of glycogen with a brief vortex and an incubation at room temperature for 10> minutes. The solution was centrifuged at 12,000g at 4°C for 10 minutes. The

supernatant was discarded and the pellet washed twice with 80% ethanol vortexing between each wash and centrifuging at 7,500g at 4°C for 5 minutes. The supernatant was discarded and samples resuspended in a total volume of 200µL of sterile water for irrigation and heated at 65°C for 10 minutes. Samples were incubated between 30 and 60 minutes on ice. After this period, samples were quantified using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Rockford, IL, USA), and 1/10<sup>th</sup> volume of 3M Sodium Acetate, 2.5 times the volume of 100% ethanol and 5mg/mL of glycogen (Ambion, Austin, TX, USA) was added, followed by a brief vortex and incubation at -20°C for 24 hours. The mixture was centrifuged at 12,000g at 4°C for 30 minutes, and the pellet washed twice with 80% ethanol with centrifugation at 12,000g at 4°C for 10 minutes following each wash. The supernatant was discarded and using the quantification values previously obtained the RNA sample was resuspended in the desired volume of sterile water to obtain a final concentration of 2µg/mL. To aid resuspension, the RNA sample was incubated at 65°C for 10 minutes. The sample was re-quantified on the Nanodrop ND-1000 spectrophotometer to determine the final concentration of the RNA.

## **2.4 Sequencing of B-Raf**

Following RT-PCR with B-Raf specific primers (Section 5.11) sequencing of exon 15 was conducted using 1µL of the exon 15 sequencing primer 5-GTGAGGGCTCCAGCTTGTATC-3 (3.3pmol/µL), 1µL of the B-Raf PCR products, 4µL BigDye3.1 kit (Applied Biosystems) and 4µL H<sub>2</sub>O. A PCR program of 95°C for 30 seconds, 50 °C for 10 seconds and 60°C for 4 minutes for 40 cycles was used. This was then cleaned with the Dye Ex 2.0 Spin kit (Quigen) using 20µL volumes. DyeEx

Kits are designed for the removal of unincorporated dye terminators directly from sequencing reactions. The resin in the spin column was vortexed for resuspension. The cap of the column was loosened a quarter of a turn to avoid a vacuum inside the spin column. At the bottom of the spin column the closure plastic was snapped off and placed in the 2mL collection tube. This was then centrifuge for 3 minutes at 101g using a Micro centaur bench-top centrifuge (Sanyo). The spin column was transfer to a clean centrifuge tube and the sequencing reaction (20µL) applied drop-wise directly onto the centre of the slanted gel-bed surface to the to the gel bed. The sequencing reaction product was collected by further centrifugation for 3 minutes at 101g using the Micro centaur bench-top centrifuge. The spin column was discarded and the eluate in the micro-centrifuge tube contained the purified DNA. Sequencing reactions were sent to the 'in house' Central Biotechnology Service (CBS) where sequence analysis was undertaken using the automated ABI PRISM 3100 Genetic Analyser.

## **2.5 One-cycle eukaryotic labelling of gene chips**

### ***2.5.1 Sample preparation***

For each gene chip a 60mm dish was used. This equated to 4 dishes for RASV12, 4 dishes for BRAF<sup>V600E</sup> and 4 dishes for RET/PTC1 since the microarray experiment was conducted in quadruplicate. I processed all the dishes for total RNA isolation and purification in an RNase-free environment (section 2.3). Following the isolation of total RNA from samples, the following procedure involving labelling of gene chips were conducted by the in-house Affymetrix service which is part of the CBS facility. This method describes the experimental procedures conducted by the CBS facility.



Total RNA was first reverse transcribed using a T7-Oligo(dT) Promoter Primer in the first- strand cDNA synthesis reaction. Following RNase H-mediated second-strand cDNA synthesis, the double stranded cDNA was purified and used as the template in the in vitro transcription (IVT) reaction. The IVT reaction was carried out with the T7 RNA polymerase and a biotinylated nucleotide ribonucleotide mix for complementary cRNA amplification and biotin labelling. These cRNA targets were then cleaned up, fragmented and hybridised to gene expression arrays. All reagents are supplied from Affymetrix (High Wycombe, UK), unless detailed otherwise.

#### *2.5.2 Quantification of RNA*

RNA was quantified by spectrophotometric analysis on the basis that 1 absorbance unit at 260nm equals 40µg/mL RNA. The absorbance was checked at 260nm and 280nm to determine RNA concentration and purity, and a A260/A280 ratio close to 2.0 indicates pure RNA. The integrity of RNA was also assessed using an Agilent 2100 Bioanalyzer, which produces an electropherogram displaying two well defined peaks which correspond to the 18S and 28S ribosomal RNAs, with a ratio of 2:1 for the 18S and 28S bands.

#### *2.5.3 One-cycle cDNA synthesis*

Poly-A RNA controls for the one-cycle cycle cDNA synthesis were first prepared. These controls provide exogenous positive controls to monitor the entire eukaryotic target labelling process. Each eukaryotic gene chip probe array contained dilutions. These were amplified and labelled with the samples, and were used to examine the

hybridisation intensities on the arrays by monitoring the labelling process independent of the starting RNA samples. For 5µg of total RNA, 2µL of the Poly-A Control Stock was diluted in 38µL of the Poly-A Control Dilution Buffer for a 1:20 dilution (first dilution). After mixing, centrifugation and the lower aqueous phase collected, 2µL was further diluted in 98µL of the Poly-A Control Dilution Buffer, producing a 1:50 dilution. This like previously, was mixed, centrifuged and the lower aqueous phase collected and 2µL was diluted in 18µL of the Poly-A Control Dilution Buffer to give a 1:10 dilution. After mixing, centrifugation and collection of the lower aqueous phase, 2ul of the final dilution was added to 5µg of total RNA sample.

#### *2.5.4 First-strand cDNA synthesis*

After thoroughly mixing the RNA samples, diluted poly-A controls and the T7-Oligo(Dt) primer, the total RNA (1µg-15µg) was added to 2ul of the diluted Poly-A RNA controls and 2µL of 50µM T7- Oligo(Dt) primer. RNase free water was finally added to a total volume of 12µL and the tube was mixed and pulsed briefly to collect the reaction mix at the bottom of the tube. The reaction mix was incubated for 10 minutes at 70°C and briefly cooled at 4°C for ~2 minutes. In a separate tube the first strand master mix was assembled (Table 2.1). This mixture was transferred to each RNAsample/T7-Oligo(Dt) primer mix, thoroughly mixed and incubated at 42°C for 2 minutes. Depending on the amount of total RNA, Superscript II was added (1µL for 1-8µg RNA or 2µL for 8.1-15µg RNA), thoroughly mixed and incubated at 42°C for 1 hour, followed by cooling the samples at 4°C for at least 2 minutes.

Component	Volume
5X 1 <sup>st</sup> Strand Reaction Mix	4µL
0.1M DTT	2µL
10mM dNTP	1µL

**Table 2.1. First-Strand Master Mix**

### *2.5.5 Second-strand cDNA synthesis*

In another tube the second-strand master mix was assembled (Table 2.2) and thoroughly mixed. To each first-strand master mix 130µL of the second strand master mix was added and incubated at 16°C for 2 hours. After which, 2µL of T4 DNA polymerase was added to each sample and incubated at 16°C for 5 minutes. Following this incubation, 10µL of 0.5M EDTA was added.

Component	Volume
RNase-free water	91µL
5X 2 <sup>nd</sup> Strand Reaction Mix	30µL
10mM dNTP	3µL
E.coli DNA ligase	1µL
E.coli DNA Polymerase I	4µL
RNase H	1µL

**Table 2.2. Second-Strand Master Mix**

### *2.5.6 Cleanup of Double-Stranded cDNA*

To the double-stranded cDNA 600μL of cDNA Binding Buffer was added and vortexed briefly. To a cDNA Cleanup Spin Column sitting in a 2mL collection tube 500μL of the sample was applied and centrifuged for 1 minute at  $\geq 8000g$ , and the flow through discarded. The remaining sample mixture was applied and centrifuged and the flow through discarded as before. The spin column was transferred to a new 2mL collection tube and 700μL of the cDNA wash buffer was added to the column, which was centrifuged for 1 minute at  $\geq 8000g$  and the flow through discarded. The spin column was then centrifuged for 5 minutes at maximum speed with the cap open and any flow through discarded as before. The spin columns were transferred into a 1.5mL collection tube, to which 14μL of cDNA elution buffer was added to the spin column membrane and incubated for 1 minute at room temperature and centrifuged for 1 minute at maximum speed for elution to occur.

### 2.5.7 Synthesis of Biotin-labelled cRNA

For this section the gene chip IVT labelling kit was utilised. The amount of cDNA to be used for each IVT reaction was calculated (Table 2.3) and transferred to RNase-free tubes.

Starting RNA Material	Volume of cDNA for IVT
10 to 100ng	12µL
1 to 8µg	12µL
8.1 to 15µg	6µL

**Table 2.3.** IVT reaction set-up

Components from the IVT labelling kit were added (Table 2.4) and centrifuged briefly and incubated at 37°C for 16 hours.

Component	Volume
Template cDNA	Variable
RNase-free water	variable
10X IVT Labelling Buffer	4µL
IVT Labelling NTP mix	12µL
IVT Labelling Enzyme Mix	4µL

**Table 2.4.** IVT labelling kit components

### ***2.5.8 Cleanup and quantification of Biotin-Labelled cRNA***

To each IVT reaction 60µL of RNase-free water was added and briefly vortexed. To this 350µL of IVT cRNA Binding Buffer was added and again briefly vortexed. 250µL of ethanol (96-100%) was then added, mixed followed by the application of 700µL of the sample to an IVT cRNA Cleanup Spin Column which is sat in a 2mL collection tube. The columns were centrifuged for 15 seconds at  $\geq 8000g$ , and the flow through and collection tube discarded. The spin columns were transferred to a new collection tube and 500µL of the IVT cRNA Wash Buffer was applied onto the spin column. The column was centrifuged for 15 seconds at  $\geq 8000g$ , and the flow through discarded. To the column 500µL of 80% (v/v) ethanol was applied and centrifuged for 15 seconds at  $\geq 8000g$ , and the flow through and collection tube discarded. The columns were then transferred to 1.5mL collection tubes and 11µL of RNase water was added and centrifuged for 1 minute at maximum speed. A further 10µL of RNase-free water was added directly to the spin column membrane and the column centrifuged 1 minute at maximum speed. Spectrophotometric analysis was then utilised to determine the cRNA yield and purity.

### ***2.5.9 Fragmentation of the cRNA for Target Preparation***

Fragmentation of the cRNA target before hybridisation onto gene chip probe arrays is crucial to obtain optimal assay sensitivity. The Fragmentation Buffer has been optimised to break down full length cRNA to 35 to 200 base fragments by metal induced hydrolysis. The fragmentation reaction mix was formulated (Table 2.5), incubated at 94°C for 35 minutes and then cooled on ice. An aliquot was saved for testing on the Agilent Bioanalyzer to check the fragmentation procedure.

Component	Volume
cRNA	20µg
5X Fragmentation Buffer	8µL
RNase-free water	Make up to 40µL

**Table 2.5.** Fragmentation Mix components

### 2.5.10 Eukaryotic Target Hybridisation

The hybridisation cocktail was mixed together (Table 2.6) and the probe arrays equilibrated at room temperature.

Component	Volume
Fragmented and labelled cRNA	20µg
Control Oligonucleotide B2 (3mM)	8µL
20X Eukaryotic Hybridisation Controls ( <i>bioB</i> , <i>bioC</i> , <i>cre</i> )	15µL
2X Hybridisation Mix	150µL
DMSO	30µL
Nuclease-free water	Make to 300µL

**Table 2.6.** Hybridisation Cocktail Mix

The hybridisation cocktail was heated to 99°C for 5 minutes and during this incubation, the array was immersed with 200µL of Pre-Hybridisation Mix through one of the septa. The array was incubated with the pre-hybridisation mix at 45°C for 10

minutes with constant rotation. The hybridisation cocktail was then heated to 45°C for 5 minutes and centrifuged at maximum speed in a microcentrifuge. The array was removed from rotation, and the pre-hybridisation mix removed and replaced with the hybridisation cocktail. This was left to hybridise to the array at 45°C for 16 hours on a rotation of 60rpm.

#### 2.5.11 Washing, staining and scanning the array

After hybridisation, the array was vented and the hybridisation cocktail removed and washed with 250µL of Wash buffer A. The automated protocol for washing and staining was performed by the Fluidics Station 450 (Table 2.7)

Step	
Post hybridisation wash #1	10 cycles of 2 mixes/cycle with Wash buffer A at 30°C
Post hybridisation wash #2	6 cycles of 15 mixes/cycle with wash buffer B at 50°C
1 <sup>st</sup> stain	5 minutes with stain cocktail 1 at 35°C
Post stain wash	10 cycles of 4 mixes/cycle with wash buffer A at 30°C
2 <sup>nd</sup> stain	5 minutes with stain cocktail 2 at 35°C
3 <sup>rd</sup> stain	5 minutes with stain cocktail 3 at 35°C
Final wash	15 cycles of 4 mixes/cycle with wash buffer A at 35°C
	Filled with Array Holding Buffer

**Table 2.7.** Automated protocol for scanning and staining



Scanning the array was performed by the GeneArray Scanner 3000 at a pixel value of 3µm and a wavelength of 570nm. This was followed by quality control analysis also conducted by CBS.

## **2.6 Invasion Assay**

Growth factor reduced BD Matrigel™ Matrix (BD Biosciences, Bedford, UK) containing reduced amounts of EGF (<0.5 ng/mL), IGF-1 (5ng/mL), PDGF (<5 ng/mL) and TGFβ (1.7ng/mL) was thawed overnight at 4°C. The stock Matrigel™ was diluted in 1%FCS media at a concentration of 500 µg/mL. In a sterile 24 plate, BD Biocoat™ Matrigel™ invasion chambers (BD Biosciences, Bedford, UK) were placed and 100µL of the diluted Matrigel™ was plated into each chamber. The plate was incubated in an oven for 4 hours at 37°C. Each chamber was then rehydrated using 100µL of 1%FCS cell culture media and incubated for 1 hour. Cells to be used in the assay were trypsinised and counted using a haemocytometer. A cell concentration of  $3 \times 10^5$ /mL was obtained and cells resuspended in media containing 1%FCS. To each invasion chamber  $3 \times 10^5$  cells were plated and 900µL of 1%FCS cell culture media was placed in the outer section of the invasion chamber. The plate was incubated at 37°C, 5% CO<sub>2</sub> for the required length of time (24 hours or 48 hours). After the invasion period, each chamber was removed and the inner cleaned with tissue. Any invaded cells which would be present on the underside of the PET membrane of the invasion chamber were fixed in methanol/acetic acid (3:1) for 10 minutes, left to air dry and stained with GIEMSA (VWR International, Poole, UK) for 30 minutes. Stained inserts were photographed at x20 magnification and 5 fields per

insert counted, with the average cell number determined. This was subsequently converted to percentage invasion.

# **3.0 Results**

### 3.1 The status of BRAF<sup>V600E</sup> mutation in thyroid cell lines

Alternative splicing of B-raf gives rise to multiple isoforms. In mice, alternative splicing encodes at least 10 isoforms of B-Raf with tissue-specific expression. The isoforms arise from the alternative splicing of exon 10 (9b), exon 8b and exons 1/2. The presence of exon 10 increases the basal kinase activity and affinity towards MEK, and exon 8b decreases kinase activity, the mechanisms of which remain unknown (Hmitou et al. 2007). Exons 1/2 promote dimerisation with Raf-1 providing an additional level of kinase activity. This allows for the modulation of kinase activity which may relate to its biological activity.

The T1799A transversion mutation induces the V600E mutation of B-Raf. This mutation accounts for ~90% of all B-Raf mutations in thyroid cancers and renders B-Raf constitutively active (Garnett and Marais 2004). Transfection of BRAF<sup>V600E</sup> in NIH3T3 cells induces transformation with constitutive ERK activation (Davies et al. 2002). The T1799A mutation present in exon 15 was analysed in the full length and a shorter isoform (excluding exon 10), and this was conducted in a number of cell lines derived from thyroid tumours.

I aimed to characterise the alternative splicing mechanism in my panel of thyroid cancer cell lines. In addition, I also aimed to determine the mutational status, an important consideration for future experiments.

### **3.1.1 Objectives**

1. To define the mutational status of B-Raf in the panel of cell lines
2. To identify B-Raf isoforms in a panel of thyroid cell lines

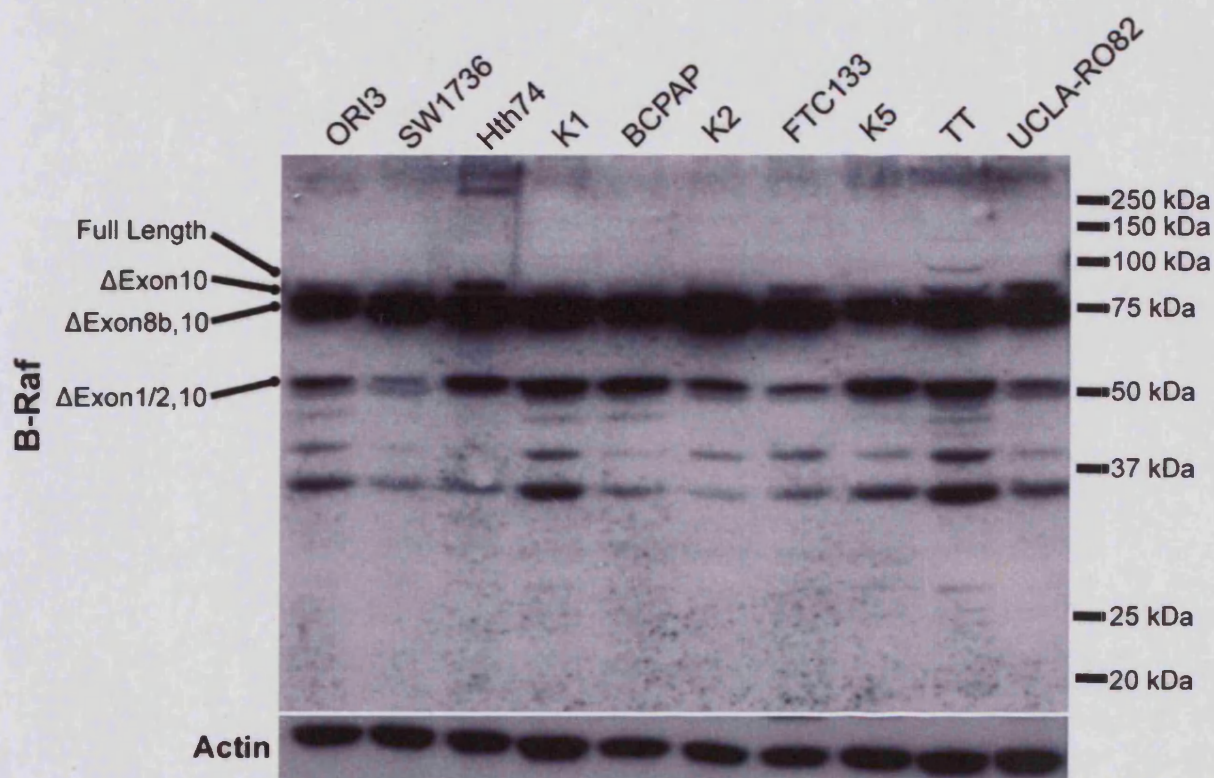
### **3.1.2 Multiple BRAF isoforms are detected in thyroid cell lines**

The regulation of kinase activity by B-Raf is complex, and alternative splicing of B-Raf is known to occur giving rise to multiple isoforms (Hmitou et al., 2007, Barnier et al., 1995). An important consideration for future experiments is establishing the B-Raf mutational status in thyroid cancer cell lines (Table 3.0). To conduct this investigation, a number of cell lines derived from various thyroid tumours were analysed by western blot analysis. The B-Raf antibody which I used was raised against a KLH-conjugated synthetic peptide corresponding to 754-765 amino acids of the human B-Raf with an N-terminal cysteine. This sequence is common to all B-Raf isoforms, and so the antibody is able to detect the full length recombinant B-Raf and B-Raf splice isoforms.

In the thyroid cancer cell lines SW1736, Hth74, FTC133, TT and UCLA-RO82 there are multiple bands detected, indicating that multiple isoforms of B-Raf are present. Predominantly detected was a shorter 74kDa isoform which corresponded to  $\Delta$ Exon10. Interestingly, in all the cell lines the full length isoform is expressed at a significantly lower level than the shorter isoforms of B-Raf indicating the dominant presence of a shorter isoform of B-Raf (Figure 6).

Cell Line	Mutational Status	Origin
ORI3	None	SV40 Transformed, immortalized
SW1736	BRAF <sup>V600E</sup> +/-	Anaplastic Carcinoma
Hth74	Unknown	Anaplastic Carcinoma
BCPAP	BRAF <sup>V600E</sup> +/-	Papillary Carcinoma
K1	BRAF <sup>V600E</sup> +/-	Papillary Carcinoma
FTC133	Unknown	Follicular Carcinoma
K5	Unknown	Papillary Carcinoma
TT	None	Medullary Carcinoma
UCLA-RO82	Unknown	Follicular Carcinoma (metastasis)

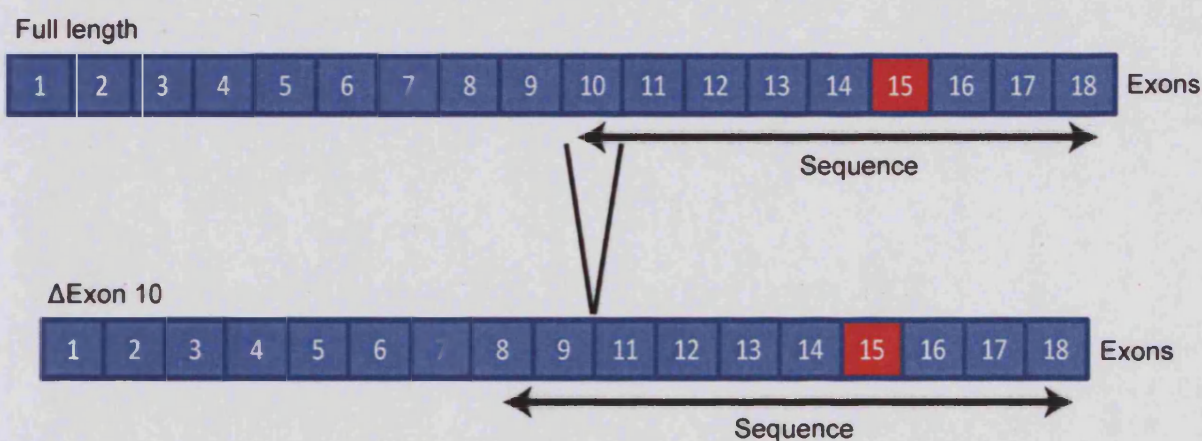
**Table 3.0.** Thyroid cell lines with associated origin and mutational status. +/- represents heterozygous BRAF<sup>V600E</sup> mutation and +/+ represents homozygous BRAF<sup>V600E</sup> mutation.



**Figure 6.** B-Raf isoforms identified by Western blot. Thyroid cell lines were probed with a B-Raf specific antibody detecting different isoforms of B-Raf. The lower molecular weight bands are deemed to be too small to be a B-Raf isoform and maybe either a non-specific artifact or due to protein degradation of the sample. To observe loading the western blot was stripped and re-probed for Actin.

### 3.1.3 The BRAF<sup>V600E</sup> mutation is detected in the BCPAP and K1 papillary carcinoma cell lines.

A number of cell lines were selected to analyse the oncogenic mutation in the transcribed functional form of B-Raf. Using RT-PCR, cDNA was generated from mRNA extracted from the cell lines, thereby eliminating genomic non-coding regions of B-Raf. Exon 10 is present in the wild-type isoform of B-Raf and is spliced out of the shorter isoforms. Amplifying the region of exons 10-18 isolated the wild-type isoform since the forward primer is specific for exon 10 and would only anneal to the cDNA of the wild-type B-Raf isoform (Figure 7). For the detection of a shorter isoform of B-Raf, RT-PCR was conducted on the region between exons 8-18. These two RT-PCR reactions were sequenced to determine the presence of the BRAF<sup>V600E</sup> mutation in exon 15 of the full length isoform (exons 10-18) and shorter isoform (exons 8-18).

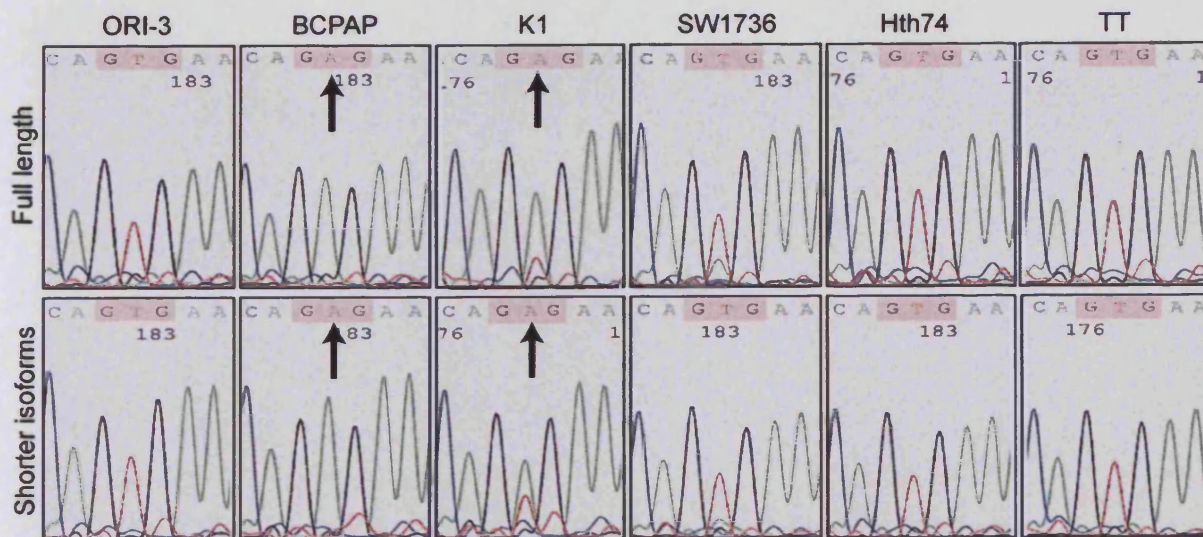


**Figure 7.** Methodology of RT-PCR sequencing for the full length and shorter isoforms of B-Raf.



The normal epithelial cell line ORI-3, displayed a red trace indicating a Thymine at the site of the point mutation confirming the presence of a valine codon (GTG), and therefore no BRAF<sup>V600E</sup> mutation. This was consistent for both the full length and short isoforms of B-Raf. The SW1736, Hth74 and TT cell lines displayed similar traces, highlighting the lack of the BRAF<sup>V600E</sup> mutation. Based on literature, the SW1736 cell line should contain a heterozygous BRAF<sup>V600E</sup> mutation. Since no mutation was detected, this indicates that the lab stock of this cell line is compromised. The mutation of B-Raf can either be present on both maternal and paternal alleles (homozygous) or on a single allele (heterozygous). The BCPAP papillary carcinoma cell line harbours a homozygous mutation in the full length isoform of B-Raf, indicated by a green trace which designates an Adenine base, and a Glutamic Acid codon (GAG). A heterozygous mutation in the full length B-Raf isoform was demonstrated in K1 indicated by the overlapping Thymine and Adenine bases at the point mutation. This demonstrates that only one B-Raf allele possesses the mutation in the full length isoform. The results were also consistent in the amplified region of exon 8-18 which also harboured the point mutation (Figure 8).





**Figure 8.** Sequencing data of selected cell lines observing BRAF<sup>V600E</sup> mutation. Region containing the mutated codon is highlighted in pink. GTG represents wild-type B-Raf codon (Valine). GAG represents mutated codon (Glutamic Acid). Black arrows highlight the presence of an Adenine base highlighting the BRAF<sup>V600E</sup> mutation.

### 3.1.4 Summary

Alternative splicing of B-raf gives rise to multiple isoforms. In mice, alternative splicing encodes at least 10 isoforms of B-Raf with tissue-specific expression. The isoforms arise from the alternative splicing of exon 10 (9b), exon 8b and exons 1/2. The presence of exon 10 increases the basal kinase activity and affinity towards MEK, and exon 8b decreases kinase activity (Hmitou et al. 2007). Exons 1/2 promote dimerisation with Raf-1 providing an additional level of kinase activity. These spliced exons modulate the kinase activity of B-Raf to mediate different biological effects (Barnier et al. 1995).

Signalling of the MAPK pathway through B-Raf induces sustained ERK activation and cellular differentiation by phosphorylation of dual-specificity kinases MEK1 and 2 which phosphorylate ERK1 and 2. The phosphorylation events shuttles ERK to the nucleus which induces the multitude of biological effects. Occurring in the cytosol, an additional level of activity is provided by the heterodimerisation of B-Raf with Raf-1, in a Ras dependent manner (Weber et al. 2001; Garnett et al. 2005). In this instance the activation of Raf-1 requires activation region and S621 phosphorylation. When B-Raf becomes activated, Raf-1 may bind in a 14-3-3 dependent manner, which activates Raf-1 to signal to MEK1/2. The dimerisation may provide B-Raf with subtle modulation of its kinase activity, which may relate to the biological outcome.

Analysis of B-Raf isoforms in my panel of thyroid cell lines revealed multiple isoforms. There was a prominent shorter isoform which based on current literature I determine to be the  $\Delta$ Exon 10 isoform. There is however a discrepancy in size of B-Raf isoforms which I identified and the isoforms identified in the literature. The study by Barnier et al. was conducted in murine cells, and this may not fully recapitulate

the situation in human thyroid. In addition, modifications such as phosphorylation may be a probable cause of why the sizes do not match.

The T1799A transversion induces the BRAF<sup>V600E</sup> mutation. This accounts for ~90% of all B-Raf mutations in thyroid cancers (Hmitou et al. 2007). BRAF<sup>V600E</sup> mimics T599/S602 phosphorylation rendering B-Raf constitutively active. Ectopic expression of BRAF<sup>V600E</sup> in NIH3T3 cells transforms cells with a 667-fold increase in kinase activity compared to wild-type B-Raf and induces constitutive ERK activation (Davies et al. 2002). In PCCL3 rat thyroid cells, BRAF<sup>V600E</sup> induces DNA synthesis, apoptosis, dedifferentiation and chromosomal instability (Mitsutake et al. 2005). In contrast, transformation by BRAF<sup>V600E</sup> with maintenance of proliferation was demonstrated in human papillary carcinoma cell lines (Liu et al. 2007). The difference in species and experimental conditions may provide an indication as to why BRAF<sup>V600E</sup> has a different effect in these studies. The BRAF<sup>V600E</sup> mutation is associated with aggressive characteristics including lymph node metastasis, extrathyroidal invasion and recurrence (Namba et al. 2003; Palona et al. 2006). Recent evidence has emerged suggesting that a homozygous mutation is correlated with increased aggression in thyroid cancer (Baitei et al. 2009). The aggressive nature of a homozygous mutation is highlighted by the BCPAP cell line and this was also shown to invade to a high degree *in vitro*.

In addition, it was shown that the phosphorylation sites of T599 and S602, where the BRAF<sup>V600E</sup> mutation occurs were not present in spliced variants. Instead these variants did not possess the N-terminal auto-inhibitory domain and only contained the C-terminal kinase domain resulting in constitutive activation. This indicates the shorter isoforms of B-Raf do not contain the V600E mutation and the mutation is restricted to the full length isoform only (Baitei et al. 2009). My investigation identified

mutations in all isoforms of B-Raf which is in contrast to the Baitei et al study. In the authors study it appears the variant isoforms which were detected were not the same ones I identified in my investigations. Further work is required to analyse if the auto-inhibitory domain is lacking in the variant isoforms which I identified, and if the mechanism of constitutive B-Raf activation proposed by Baitei et al. is applicable to the thyroid cell lines I used.

This investigation was conducted to characterise B-Raf alternative splicing and the status of BRAF<sup>V600E</sup>. For future experiments, it was important to check if the full length isoform of B-Raf was the correct one to use for the *in vitro* model of PTC. I therefore aimed to select an appropriate isoform of B-Raf containing BRAF<sup>V600E</sup>. The full length isoform was not the most prevalent; however, the full length isoform contains exon 10 and 8b to modulate the basal kinase activity and affinity towards MEK and exons 1/2 which are required for heterodimerisation with Raf-1.

This means the chosen B-Raf isoform has the full array of biological activity. The presence of the mutation is not a factor for selection since the mutation is present in all isoforms. For the production of the BRAF<sup>V600E</sup> PTC model the full length isoform of BRAF<sup>V600E</sup> was utilised in a retroviral vector.

## **3.2 Microarray Expression Gene Profiling**

To identify candidate genes involved in invasive behaviour in thyroid cancer I utilised the Microarray platform to achieve this aim. Young human primary thyrocytes infected with the respective oncogenes (RASV12, BRAF<sup>V600E</sup> and RET/PTC1) were genetically profiled in quadruplicate on an Affymetrix microarray platform using U133 PLUS 2.0 gene chips by the in-house Affymetrix service. This particular array contains 47400 transcripts, totalling 38500 genes and 54000 probe sets allowing for comprehensive analysis of genome-wide expression on a single array. The primary aim for this experiment was to identify potential candidate genes regulating the invasive phenotype of thyroid cancer. The microarray study was also used to observe the genetic profiles and highlight the molecular mechanisms the each oncogene infected thyrocytes, providing further insight into the biology of FA and PTC. The microarray study was subjected to statistical analyses conducted by Dr. Peter Giles who specialises in bioinformatic analysis and is a member of the in-house Affymetrix array service.

### **3.2.1 Objectives**

1. Prepare oncogene infected thyrocytes for Microarray gene profiling
2. Conduct Microarray gene profiling quality control metrics
3. Analyse Microarray datasets to identify markers of tumour invasion
4. Identify candidate genes from literature analysis
5. Characterise invasive ability of oncogene expressing thyrocytes

### **3.2.2 Preparation of oncogene infected thyrocytes**

As detailed in the materials and methods, the *in vitro* models of RASV12, BRAF<sup>V600E</sup> and RET/PTC1 infected thyrocytes were produced by plating primary thyrocytes derived from a single patient into 12x60mm dishes (4 dishes per model). This effectively provided me with a model for FA (RASV12) and two models for PTC (BRAF<sup>V600E</sup> and RET/PTC1). Total RNA was extracted and the one-cycle eukaryotic labelling of gene chips followed by quality control metrics was conducted. Each model was conducted in quadruplicate to give the microarray experiment adequate power.

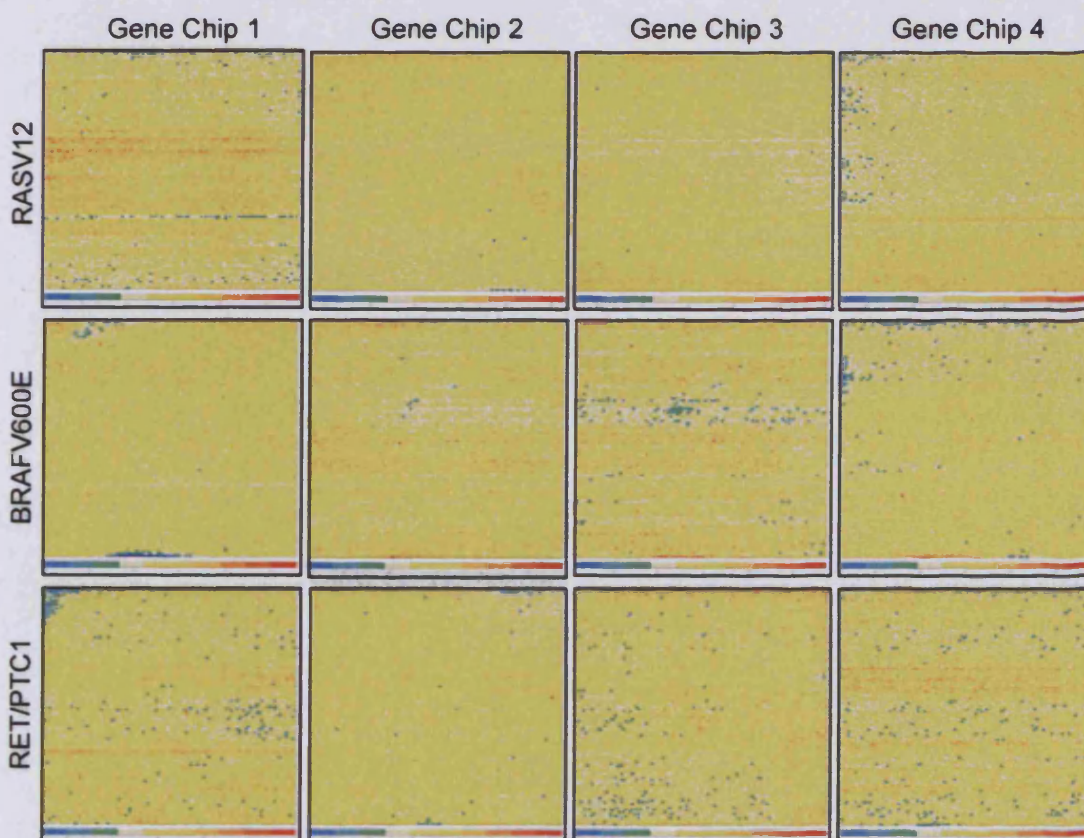
### **3.2.3 Microarray quality control metrics**

Microarray technology is a rapidly advancing, expensive yet informative biological technology. There are a number of variables which require careful consideration, and may affect the result, precision and accuracy. The starting sample material is crucial, as samples may display variable RNA quality. The use of quality control metrics assures accurate replication within sample sets and technical variation within the experiment ensuring consistency with accurate and precise gene expression data.



### 3.2.3.1 Spatial Artefacts

To highlight physical imperfections and defective hybridisation within each gene chip, the gene chips were pseudo coloured to observe signal intensity. This quality control metric contains a localised scaling factor which is defined as a multiplication factor applied to every signal value in each chip. Scaling factors should be consistent within a sample set, highlighted by a prominent yellow colour. Poor quality is highlighted by indicated by extreme colours (blue/red). This forms the basis for the detection of hybridisation and physical chip defects. In this Microarray experiment, a prominent yellow colour with consistent scaling factors was observed for the gene chips indicating no physical imperfections or defective hybridisation (Figure 9).

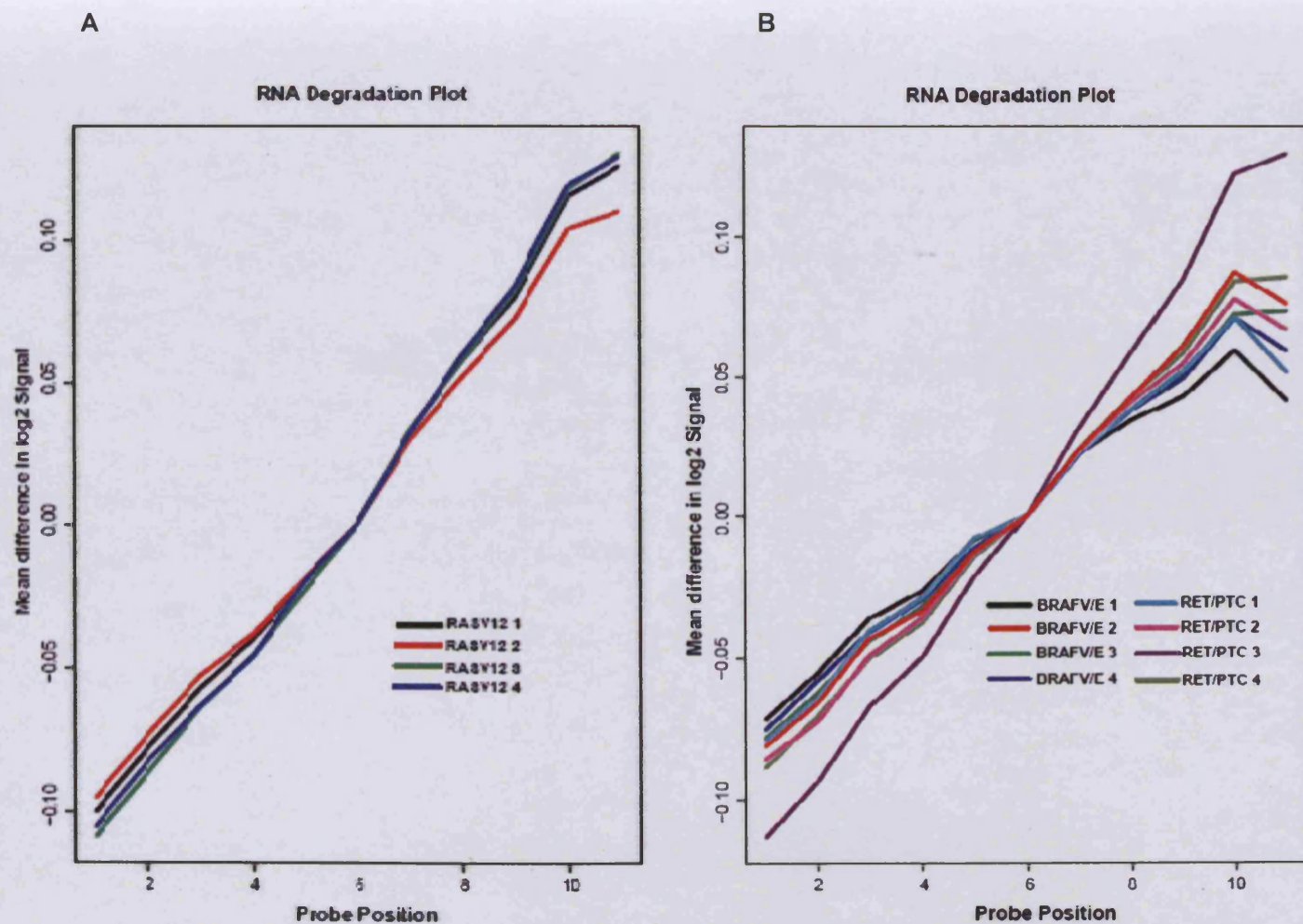


**Figure 9.** Spatial artifact quality control analysis. Gene chips used in the microarray study were pseudo-coloured to observe signal intensity for the detection of physical chip defects. The scaling factor is indicated below each gene chip.

### 3.2.3.2 RNA degradation

RNA degradation plots observe the difference in the quality of RNA for each replicate. The degradation of RNA initiates at the 5' end and terminates at the 3' end of the RNA molecule, and the plots display the expression of the 5' to 3' position of probes, random degradation and polyA-primed synthesis. The average intensity of the probes is classified in the order of their proximity to the 5' end of the gene for each gene chip. Each line displayed corresponds to an array, and the slope indicates the potential RNA degradation of the material hybridised to the array. In general, the greater the slope the more degraded the sample and overall decreased quality of data (Figure 10), but consistency is the most important factor. A single gene chip, RET/PTC1 chip 3 displayed a higher trend indicating increased degradation and poor quality of RNA, and therefore this particular chip was excluded from further analysis.





**Figure 10.** RNA Degradation plots of RASV12, BRAF<sup>V600E</sup> and RET/PTC1 genechips. **A**, RASV12 gene chips. **B**, BRAF<sup>V600E</sup> and RET/PTC1 gene chips. The average intensity of the probes were classified in the order of their proximity to the 5' end of the gene for each gene chip. Each line displayed corresponds to each gene chip with the slope indicating the potential RNA degradation of the material hybridised to the array.

### 3.2.3.3 *Percent present genes and internal control genes*

Percent present genes is defined as a fraction of probe sets 'expressed' in relation to the entire array, as produced by the Affymetrix MAS 5.0 statistical algorithm. This algorithm works by taking a set of Perfect Match (PM) and Mis-Match (MM) probes, and using these to generate a single value representing the estimated amount of transcript. To do this, files containing array images were processed to produce measured intensities for each probe on the array. The values depend on multiple factors including cell/tissue type, biological or environmental stimuli, probe array type and overall RNA quality. Logically, replicate arrays should have similar values, and in the experiment all gene chips had similar values (Figure 11). The absolute % present values were also similar between the gene chips.

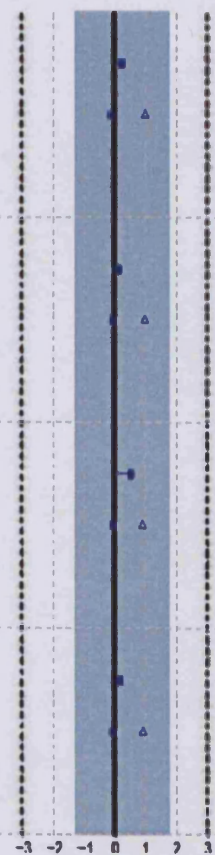
To assess sample and assay quality the internal control genes GAPDH and  $\beta$ -actin, which are ubiquitously expressed were used. The signal values of the 3' probe sets for GAPDH and probe sets were compared with signal values of the corresponding 5' probe sets. Since the gene chips contain separate probe sets against the 3', mid and 5' regions of the GAPDH transcript, a high ratio may indicate presence of truncated transcripts which may occur due to RNA degradation. Both GAPDH 3'-5' (circles) and  $\beta$ -actin (triangles) were plotted for each gene chip. The GAPDH ratios were  $<1$  which is the suggested maximum value of 1.25. The  $\beta$ -actin ratios were well within the limit of 3 (as stated by Affymetrix) for all the gene chips (Figure 11).

A

▲ AFX-HSAC07.X00351.3/8  
● AFX-HUNGAPDH.M33197

QC Stats

44.35%  
RASV12 4  
52.21  
45.28%  
RASV12 3  
48.72  
43.02%  
RASV12 2  
47.41  
45.74%  
RASV12 1  
49.50

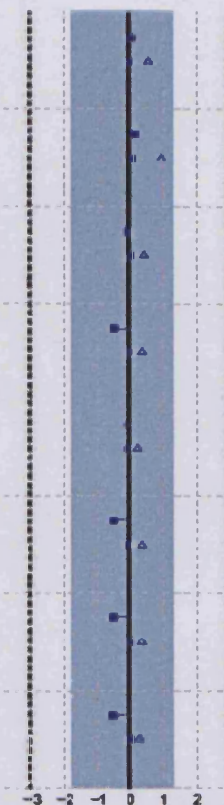


B

▲ AFX-HSAC07.X00351.3/8  
● AFX-HUNGAPDH.M33197

QC Stats

41.68%  
RET/PTC 4  
50.36  
43.44%  
RET/PTC 3  
47.73  
43.38%  
RET/PTC 2  
48.17  
43.33%  
RET/PTC 1  
51.00  
43.88%  
BRAFV6 2  
47.41  
44.52%  
BRAFV6 3  
50.92  
44.83%  
BRAFV6 2  
53.90  
45.68%  
BRAFV6 1  
40.17



**Figure 11.** Percent present genes and internal control genes analysis of RASV12, BRAF<sup>V600E</sup> and RET/PTC1 genechips. **A**, RASV12 and **B**, BRAF<sup>V600E</sup> and RET/PTC1 gene chips. The percent present genes are highlighted in blue. The GAPDH internal control gene ratio is highlighted by the blue circles, and the  $\beta$ -actin ratio is highlighted by the blue triangles.

#### 3.2.3.4 QC Summary

Following the QC metrics, only the RET/PTC1 gene chip 3 revealed increased RNA degradation which would affect its hybridisation and display a variability within the sample dataset, and so was subsequently removed from further analysis. All the other gene chips displayed consistency quality control analyses with percent present and internal control genes all within the desirable limits.

### 3.2.4 Microarray Analysis Design

As a general rule, microarray data can be analysed in 2 different ways. (1) Analysing the data to find the answer to a set hypothesis. (2) Analysing the data in order to find interesting biological information within the dataset. Using both principles, I analysed the datasets for my microarray experiment to obtain a thorough analysis. This will allow me to identify novel genetic markers and to observe any other relevant biological information associated with the oncogenic transformation of RASV12, BRAF<sup>V600E</sup> and RET/PTC1. To conduct this, the experiment was analysed using 2 separate statistical tests; a *t-test* to determine the hypothesis and an Analysis of Variance (ANOVA) test to identify relevant biological information.

### 3.2.5 Gene Profiling Analyses

I chose to use the U133 PLUS2.0 microarray gene chip array since it offers complete coverage of the Human Genome U133 Set plus 6,500 additional genes for analysis of over 47,000 transcripts. This array offers the best solution for this investigation. It is also the first time that this array has been used with oncogene infected thyrocytes, as the genetic backgrounds of the oncogenic conditions have not previously been studied in this manner.



### 3.2.5.1 Statistical *t*-test analysis of RASV12, BRAF<sup>V600E</sup> and RET/PTC1.

The RASV12 oncogene is present in most benign non-invasive FA. In contrast, BRAF<sup>V600E</sup> and RET/PTC1 are both present most in invasive PTC, including occult micro PTC. To highlight specific genes involved with invasive behaviour of PTC, the BRAF<sup>V600E</sup> and RET/PTC1 datasets were combined and compared with the RASV12 dataset using a *t*-test. This test is designed to identify the genes which are statistically differentially expressed between the invasive and non-invasive models. The method of Benjamini and Hochberg (1995) *False Discovery Rate* (FDR) was conducted. This method correct for multiple testing and false discovery rates among the significant values obtained ( $p=0.05$ , 95% chance of differential gene expression and 5% chance of a false positive). In this experiment the FDR was set to 0.1% for the *t*-test, which significantly reduces the amount of false positives within the data. The data was subjected to hierarchical clustering, using the Euclidean distance measure. By definition, the Euclidean distance measures the absolute distance between two genes and can be visualised by a dendrogram, where the shorter the distance the greater the similarity. To visualise this concept, heat maps were produced using a colour coded matrix, where red signifies high expression, green low expression and the median expression as black (created using logs transformed median centred data). The heat maps observing the significant differential gene expression were produced using the R Package For Statistical Computing. This complete method of applying FDR, hierarchical clustering and generating heat maps were applied to the other statistical tests. For the *t*-test a probe set of 705 differentially expressed genes was identified (FDR=0.1%), displaying 2 unique clusters (Appendix Figure 1). Cluster 1 contained 334 up-regulated genes in RASV12, with genes either down-regulated or at median expression in both

BRAF<sup>V600E</sup> and RET/PTC1. Cluster 2 contains 334 down-regulated genes in RASV12 with genes up-regulated or at median expression in BRAF<sup>V600E</sup> and RET/PTC1 conditions.

To analyse the clusters in further detail, the primary biological functions were analysed. This would provide an insight into any processes which are present in the oncogenic models with correlation with malignancy. In order to gain an insight into the key biological information in each cluster and thereby present key information pertaining to invasion and tumourigenesis in the oncogenic conditions, over-representation analysis was conducted using the Metacore™ analysis suite (GeneGo; <http://portal.genego.com>). The package as a whole offers the storage and annotation of probe sets incorporating high quality curated databases of human protein-protein and protein-DNA interactions, transcription factors, signalling and metabolic pathways and effects of bioactive molecules. The incorporation of these databases facilitates detailed genetic background of a single gene together with its interaction with other genes and pathways. Within the software package, the GeneGo pathway map folder enrichment tool was used. This is a collection of manually created pathway maps, grouped hierarchically into folders according to main biological processes in the interrogated gene list. Interrogating cluster 1 (Appendix Figure 2) and cluster 2 (Appendix Figure 3) respectively revealed a distinct set of biological functions elucidating the key phenotypes in each of the clusters. By default, this analysis tool specifies an FDR of 0.05 which lists the significant processes containing no more than 5% of false positive results. The orange bars representing experiment mappings that do not pass through this level are light orange/opaque, whilst the dark orange bars are deemed significant (Appendix Figure 2 and 3).

The 334 differentially regulated probe sets (Appendix Table 1.0) represented gene transcripts from a diverse range of functional categories that regulate a multitude of biological functions. Cluster 1 had an expression pattern of up-regulation in RASV12 and median/down-regulation in BRAF<sup>V600E</sup> and RET/PTC1. The map folder tool identified vascular development (angiogenesis) as the most significant process (Appendix Figure 2). Further analysis revealed the upregulation of the VEGFA, a potent mediator of angiogenesis. The expression of FN was found expressed highly in the RASV12 model. FN is associated with EMT and is utilised as a differential marker of thyroid malignancy. The identification of FN and VEGFA is indicative of a malignant phenotype in this particular cluster (Appendix Figure 2).

The analysis also highlighted an EMT response as 'Regulation of epithelial-to-mesenchymal transition (EMT)', 'TGF-beta-dependent induction of EMT via MAPK' and 'TGF-beta-dependent induction of EMT via RhoA, PI3K and ILK' were identified. RhoA expression was highly up-regulated in the RASV12 oncogenic condition. The upregulation of this gene in RASV12 implies a migratory phenotype. Rho regulates the rear of the cell by stimulating the formation of stress fibers critical for actin-myosin mediated cell contraction (Wu et al. 2004). This forms the mechanism together with other GTPases to induce cellular movement.

The enrichment analysis identified an inflammatory response in the RASV12 model with the Cell adhesion-Chemokines and adhesion pathway highlighted. The genes associated with this pathway and upregulated in RASV12 included CD44, interleukin-8, VEGFA, FN, syndecan-2 and caveolin-2 upregulation. The VEGFA and FN genes have been previously associated with other pathways; however, genes can participate in multiple pathways thus explaining why genes are identified more than once.



The analysis also identified the 'Tissue remodeling and wound repair' process as a significant hit. The pathways identified in this process included 'Cell adhesion-ECM remodeling', 'Cell adhesion\_Alpha-4 integrins in cell migration and adhesion' and 'Development-EGFR signaling via PIP3'. The genes which I found interesting with this process include the EGF receptor and PTEN tumour suppressor protein, which were both up-regulated in RASV12. Integrins  $\alpha 4$  and  $\beta 1$  were also up-regulated in RASV12, and these are essential for cell migration and polarity in certain cell types. PTEN is a negative regulator of the Akt pathway, and plays an important role in cell cycle arrest and apoptosis. RASV12 transformation results in increased activity of Akt and subsequently a natural increase in PTEN expression. The PI3K pathway is one of the effector pathways of RASV12 transformation, which explains the increase of Akt and PTEN in my analysis. The increase in PTEN may be an apparent cellular mechanism to counter the effects of increased Akt signalling.

The final significant process identified in cluster 1 was apoptosis, and within this process 'Apoptosis and survival-Endoplasmic reticulum stress response pathway' was identified. A specific gene relating to this pathway was the heat shock protein 70 which was up-regulated in RASV12. The upregulation of this gene is linked to stress stimuli/signaling, and the MAPK pathway is central to the regulation of heat shock protein 70, and therefore is in concordance with the MAPK signalling cascade which RASV12 functions in.

Cluster 2 had 334 genes which were down-regulated in the RASV12 model and up-regulated in BRAF<sup>V600E</sup> and RET/PTC1 model (Appendix Table 2.0). The map folder enrichment of cluster 2 identified tissue remodelling and wound repair as the most significant (Appendix Figure 5). In this process, the role of ACM3 and ACM4 in keratinocyte migration was identified. Keratinocyte migration plays a key role in re-

epithelialisation of skin wounds. Central mediators of this process are the muscarinic cholinergic receptors (ACM).

The calcium signalling process was also enriched in cluster 2, and the calmodulin genes are important in this process. These genes act as mediators of many enzyme and calcium ion channel function. Calcium can regulate the actomyosin-based contractile forces and the formation and disassembly of cell-matrix adhesions. Calcium sensitive actin binding proteins regulate the structure and dynamic behaviour of the cytoskeleton and cell migration (Lee et al. 1999). In addition, calcium is essential for thyroid hormone synthesis. TSH produced by the anterior pituitary gland fine-tunes calcium signals in the thyroid.

The second enriched process identified in cluster 2 was apoptosis. Within this process the 'DNA damage-Role of SUMO in p53 regulation pathway' was highlighted. The genes regulating this pathway include Akt1, PML and SAE1. p53 acts as an inducible, sequence-specific transcription factor on genes whose products regulate cell-cycle progression and apoptosis. PML functions as a transcription factor and tumour suppressor and its expression is cell-cycle related and functions to regulate the p53 response to oncogenic signals. Its upregulation in BRAF<sup>V600E</sup> and RET/PTC1 affects the expression of p53 gene expression. Akt1 upregulation may be a contributing factor inducing increased expression of PML and p53. The Akt pathway may also be attributed to protein synthesis where oncogenic transformation induces an increase in protein synthesis, a process which was also highlighted in the cluster 2 (Bader and Vogt 2004).

The enrichment of protein degradation process contained the 'Proteolysis-Putative ubiquitin pathway', 'Development-WNT signaling pathway' and 'Degradation of beta-catenin in the absence WNT signalling pathway'.

In cluster2, immune response and inflammatory response were identified. Although separate processes, a synergistic relationship between inflammation and an immune response is common in a number of cellular contexts including tumour progression. The pathways highlighted in the inflammatory process included 'Immune response-Inhibitory action of Lipoxins on pro-inflammatory TNF-alpha signalling', 'Cell adhesion-Chemokines and adhesion', 'Immune response-CD28 signalling' and 'Immune response\_IL-6 signalling pathway. The pathways highlighted in the immune system response process included 'Immune response-IL-15 signalling via JAK-STAT cascade', 'Development-GM-CSF signalling', 'Immune response-IFN alpha/beta signalling pathway', 'Immune response-CD28 signalling' and 'Immune response-IFN gamma signalling pathway. These responses are often observed within a tumour environment where the immuno-surveillance mechanisms mediate the immune responses.

A number of processes were identified in both clusters for example, amino acid metabolism and its regulation, apoptosis, mitogenic signalling and tissue remodeling and wound repair. However, the genes regulating these processes are unique to cluster 1 and cluster 2 respectively.

### 3.2.5.2 Analysis of Variance (ANOVA) analysis

To identify any genetic signature changes between the 3 oncogenic conditions the statistical ANOVA test was utilised. This method calculates the probability of finding the observed differences in means between two or more conditions. It is important to consider the effect of multiple testing of each probe array where there is a probability of 5% of making a false positive error if the P-value is set at 0.05. The p-values obtained from the ANOVA analysis were subjected to FDR correction, which was set at 0.01% to reduce the amount of false positives within the data. The data was then subjected to hierarchical clustering, using the Euclidean distance measure. Heat maps were produced observing the differential gene expression of the probe sets for each sample. Visualisation of the heat maps was conducted using the web-based Microarray Data Review and Annotation System (MADRAS; <http://madras.uwcm.ac.uk>), developed in-house by the Microarray Bioinformatics Group at Cardiff University was utilised.

Significant changes in transcriptional modulation were observed in the RASV12, BRAF<sup>V600E</sup> and RET/PTC1 respectively. This ANOVA analysis identified 690 differentially expressed genes (FDR of 0.01%) with consistent replication and a clear differential expression pattern for each oncogenic condition (Appendix Figure 4). To digest the data, heat maps were produced and the gene list was split (using the first 5 significant bands on the dendrogram) into 5 individual clusters displaying modulation of expression between each oncogenic condition (Table 3.1).

Cluster	Number of Probe sets	RASV12	BRAF <sup>V600E</sup>	RET/PTC1
1	121	Median/low	High	Low
2	96	Median/low	Median/low	High
3	142	Low	Median	High
4	209	High	Median	Low
5	34	High	Low	Median

**Table 3.1.** Details of hierarchical clustering of genotype-specific genes in RASV12, BRAF<sup>V600E</sup> and RET/PTC1 oncogene infected thyrocytes. Columns correspond to the oncogene and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation. Table displays number of probe sets in each cluster with overall expression pattern corresponding to each oncogenic condition.

The probe sets identified for each cluster were categorised according to biological function. The 5 clusters represented differential patterns of expression with a diverse range of biological functions, analysed by the pathway map tool provided by GeneGo. The unique modulation of genes within each oncogenic condition indicates that the oncogenic event dictates a unique genetic profile. The pattern although unique show distinct similarities, and is observed when looking at the probe sets contained within each cluster and the respective integral pathway responses.

The 121 probe sets in cluster 1 had an expression pattern of low/median expression in RASV12 and RET/PTC1 with up-regulation in BRAF<sup>V600E</sup> (Appendix Figure 5). The enrichment of the biological processes contained within the cluster strongly identified tissue remodelling and wound repair. The pathways contained within this process included 'Cell adhesion-Gap junctions', 'Development-BMP signaling', 'Development-TGF-beta receptor signaling' and 'Cell adhesion-Cadherin-mediated cell adhesion'. The genes which were up-regulated in the BRAF<sup>V600E</sup> model included Plakoglobin, cingulin, ZO-3, SMURF1 and SMURF2. Plakoglobin is a component of desmosomes

and adherence junctions. This implies a role of desmosomes and/or adherence junctions in the BRAF<sup>V600E</sup> model. Also identified were the tight junction proteins, ZO-3 and cingulin implying that tight junctions are a component of BRAF<sup>V600E</sup> colonies. Taken together the identification of cell-cell contacts in the BRAF<sup>V600E</sup> model indicates epithelial architecture. It also reinforces the notion that BRAF<sup>V600E</sup> colonies maintain epithelial differentiation markers despite expressing the BRAF<sup>V600E</sup> oncogene.

The ANOVA cluster 2 containing 96 probes had only one significant process which was an immune system response (Appendix Figure 6). This cluster contained probes up-regulated in the RET/PTC1 model with down/median expression in RASV12 and BRAF<sup>V600E</sup> models. The immune system response pathways regulated by RET/PTC1 oncogene included 'Immune response-Classical complement pathway', 'Immune response-Lectin induced complement pathway', 'Immune response-Alternative complement pathway' and Immune response-Antigen presentation by MHC class II'. These significant findings highlight the initiation of the complement pathway within this cluster with the upregulation of C1S, C3 and SERPING1 genes in the RET/PTC1 model.

The ANOVA cluster 3 containing 142 genes down-regulated in the RASV12 model, median expression in the BRAF<sup>V600E</sup> model and up-regulated in the RET/PTC1 model (Appendix Figure 7). Pathways identified within this cluster included 'Cell cycle-Nucleocytoplasmic transport of CDK/Cyclins', 'Cell cycle-Regulation of G1/S transition (part 2)', 'Cell cycle-Start of DNA replication in early S phase', 'Cell cycle-Spindle assembly and chromosome separation', 'Cell cycle-Regulation of G1/S transition (part 1)' and 'Cell cycle-Role of 14-3-3 proteins in cell cycle regulation'. These pathways were highlighted under the cell cycle and its regulation process.

Other map folder processes included Cardiac Hypertrophy, Protein synthesis, DNA-damage response, Inflammatory response, Myogenesis regulation, Apoptosis and Immune system response. These pathways recapitulate the findings from the *t-test* cluster 2 (Appendix Figure 5).

Cluster 4 contained 209 probes with upregulated probes in the RASV12 model. This cluster recapitulates the processes and pathways which were identified in *t-test* cluster 1 (Appendix Figure 4). This cluster contained genes regulating 'vascular development (angiogenesis)', 'Cell differentiation', 'Tissue remodeling and wound repair', 'Mitogenic signaling' and 'Inflammatory response' (Appendix Figure 8). These were all observed in the *t-test* cluster, and shows although the test is different, the general expression pattern remains relatively consistent.

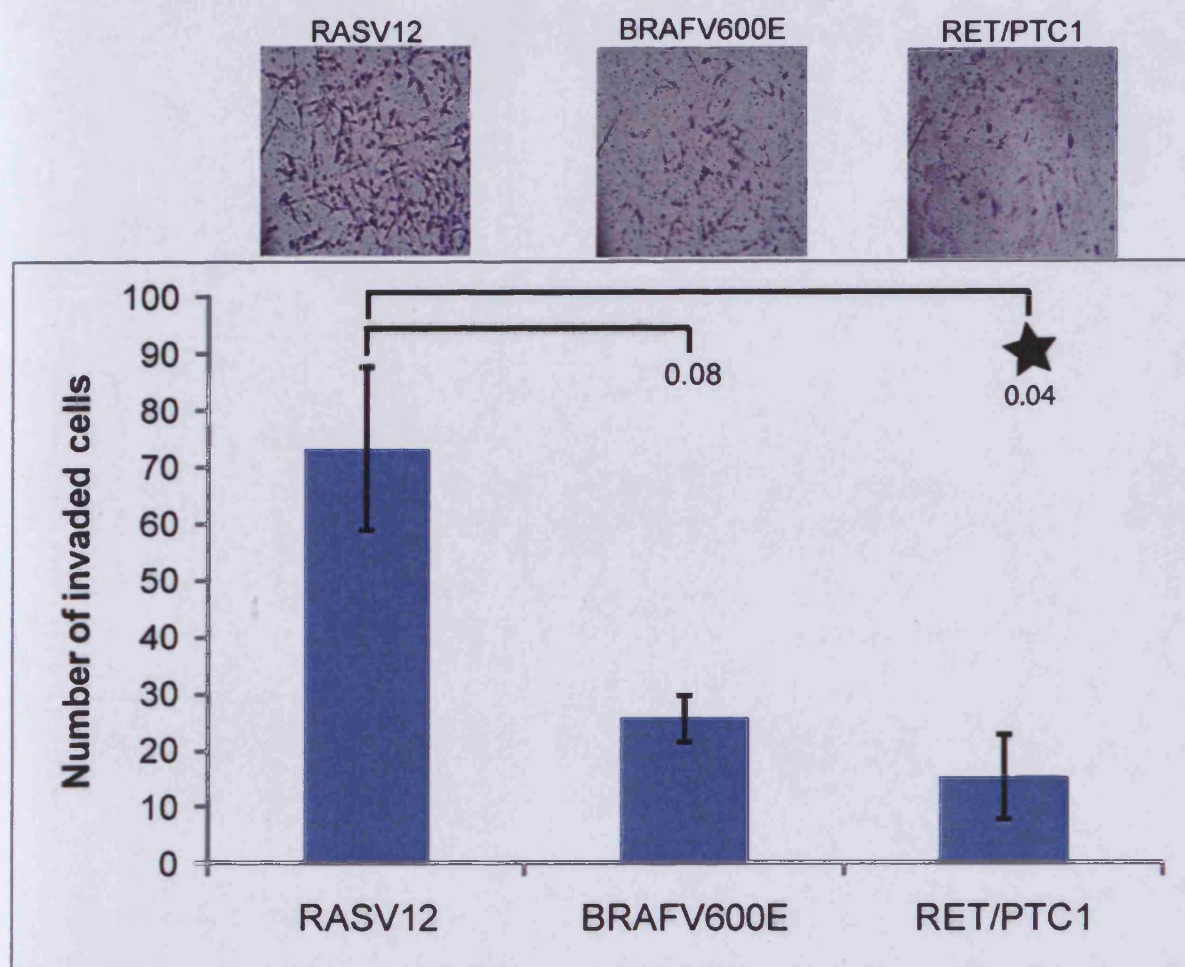
Cluster 5 containing 34 probes also displayed up-regulated probes in RASV12, did not reveal any significant map folders or individual pathways (Appendix Figure 9). The 'Apoptosis and survival-Endoplasmic reticulum stress response pathway' was identified as the highest ranked pathway (Appendix Figure 9). Only heat shock protein 70 and protein disulfide isomerase family A genes were identified. In cluster 2 of the *t-test* (Appendix Figure 5) Apoptosis was also identified. This expression pattern recapitulates these results observed in the *t-test* cluster.

### 3.2.6 Invasive behaviour of the *in vitro* models

To investigate whether microarrays could be utilised to identify candidate genes involved in invasion, I sought to characterise the invasive capability of my primary models. This would thereby link mRNA expression to the invasive behaviour of the PTC models.

To do this I had established an invasion assay using a modified boyden chamber using negative control cell lines ORI-3, MCF10A and invasive cell lines FTC236 and MDAMB231 (data not shown). In an ideal situation, the assay would be performed three times with triplicate chambers. However, there is a limited supply of primary thyroid cells. I therefore conducted a single experiment to indicate the invasive capability of the RASV12, BRAF<sup>V600E</sup> and RET/PTC1 models. The RASV12 cells were highly invasive, a finding which was unexpected, as RASV12 was assumed to be a model for non-invasive FA. Both BRAF<sup>V600E</sup> and RET/PTC1 also displayed invasive capability but at a decreased level (Figure 12). The RET/PTC1 oncogenic condition was significantly decreased when compared with the RASV12 model ( $p=0.04$ ). The decreased invasion in both BRAF<sup>V600E</sup> and RET/PTC1 may be explained by their proliferative rates. Although BRAF<sup>V600E</sup> and RET/PTC1 oncogenes display decreased proliferative rates compared with RASV12, they retain their invasive behaviour in the invasion assay. This is consistent with *in vivo* PTCs which also display a low proliferative rate.





**Figure 12.** Representative view and the number of invaded cells of the RASV12, BRAF<sup>V600E</sup> and RET/PTC1 models. Each oncogenic condition was assayed in triplicate. Invasion was conducted for 24 hours with a Matrigel concentration of 50µg/µl. After invasion, inserts were fixed and stained with GIEMSA to visualise invaded cells. Cells were counted and the average number of invaded cells calculated. *T*-test analysis was conducted by comparing RASV12 with the BRAF<sup>V600E</sup> model and RASV12 with the RET/PTC1 model. Significance between RASV12 and RET/PTC1 ( $p=0.04$ ) is highlighted by a black star.

With the discovery that RASV12 transformed cells invade Matrigel, the identification of novel candidate genes could not be achieved since there is no appropriate non-invasive control within the profiling experiment. Without comparing an invasive model with a benign model, identification of genes involved with invasion is not possible at this stage. Further microarray gene profiling of a benign model such as normal follicular epithelial cells would provide a non-invasive model to compare against.

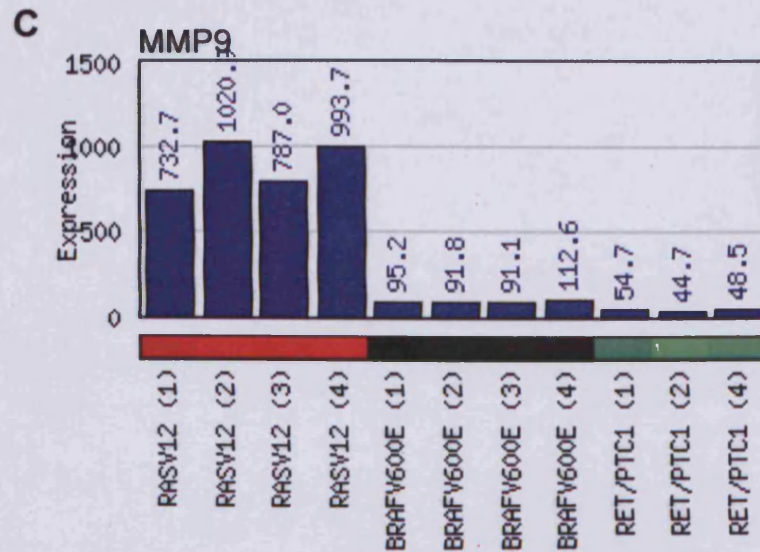
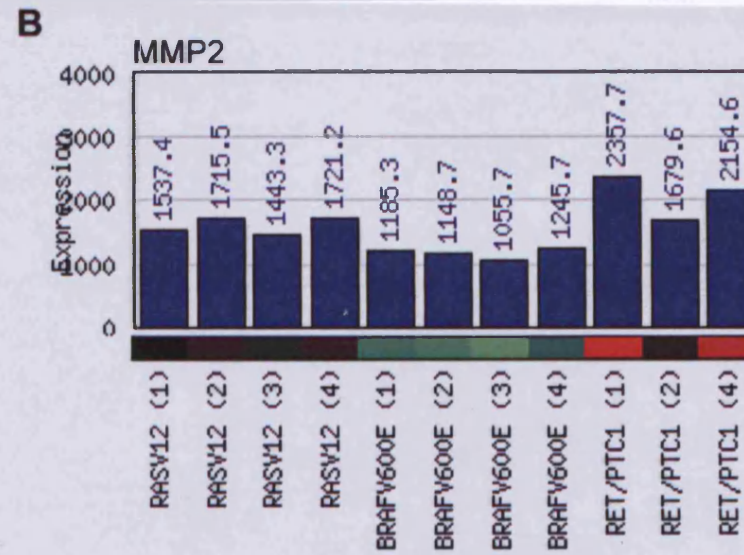
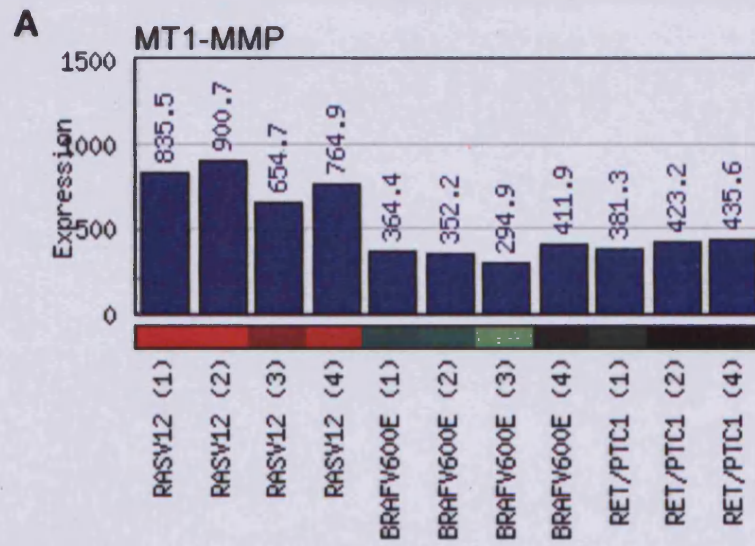
However, I conducted a literature analysis for potential candidates and profiled their expression using my microarray experiment to provide an indication that they may be involved in invasive behaviour.

### **3.2.7 Identification of candidate genes**

To further explore the genes which may regulate this invasive phenotype in these cells, an extensive literature search revealed a number of genes which mediate an invasive phenotype. In addition, many studies conduct a robust check of a number of genes to determine that datasets obtained from the analysis are both accurate and precise and are in conjunction with the biology of FA and PTC or benign vs. malignant thyroid tumours. There are a number of studies which have indicated a selection of genes which show a consistent genetic profile in benign or malignant tumours of the thyroid.

There are however, inherent factors which are taken into consideration. These are the variability in the microarray platform, methods of analysis, sample type and preparation means. The gene expression signatures are most likely to differ between studies, however a select few genes which are known to be associated with either specific tumours or malignancy. A number of studies investigating thyroid neoplasia corroborate their data with genes analysed through other experimental procedures and studies. Visualisation was conducted using a web-based Microarray Data Review and Annotation System (MADRAS) which provides expression values of individual probes for the each of the gene chips.

In the process of tumour invasion ECM degradation/remodelling is extremely important. MMPs degrade ECM products and are associated with cancer invasion. The primary MMPs which have been implicated in thyroid invasion are the membrane bound MMP (MT1-MMP), MMP-2 and MMP-9 (Hofmann et al. 1998; Buerge et al. 2009). The expression of MT1-MMP on average was up-regulated in RASV12 compared with BRAF<sup>V600E</sup> and RET/PTC1. The expression of MMP-2 was detected at median expression, whilst RET/PTC1 showed up-regulation and BRAF<sup>V600E</sup> was slightly down-regulated. The expression of MMP-9 was increased in the RASV12 model, whilst both BRAF<sup>V600E</sup> and RET/PTC1 had reduced expression (Figure 13).

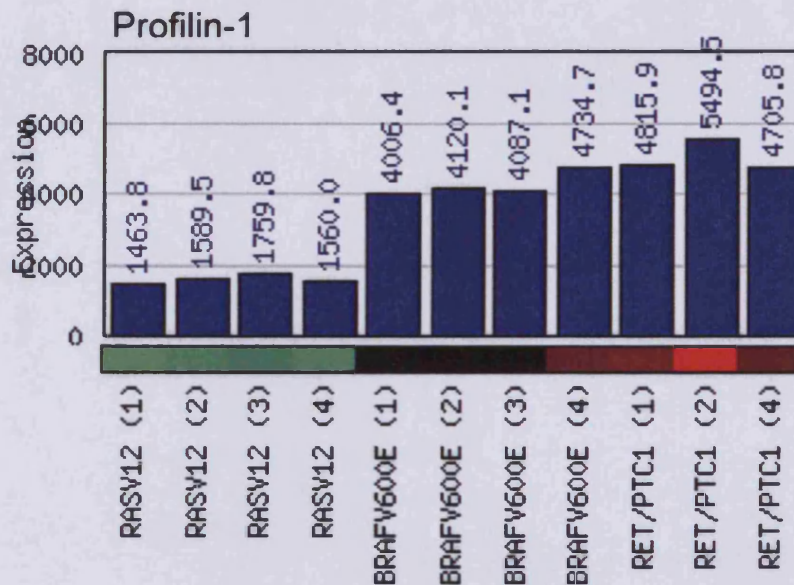


**Figure 13.** Expression profiles of Matrix Metalloproteinases determined in MADRAS. **A.** MT1-MMP **B,** MMP-2 and **C,** MMP-9. Red signifies up-regulation, green signifies down-regulation and black signifies median expression

In parallel, from the literature analysis I identified genes which have a potential role in thyroid tumourigenesis. The expression of these potential targets in my model systems were observed using MADRAS.

1) Profilins are a low molecular weight family of proteins, of which there are multiple isoforms and function to regulate the actin cytoskeleton. Profilin-1 was described initially as a G-actin binding protein, and has a primary role in sequestering actin, inhibiting actin polymerization. However, profilin can also promote actin assembly through the nucleotide ADP to ATP exchange, shuttling ATP-bound actin to the free barbed ends of actin filaments. It is therefore highly significant in mediating a motile phenotype. The polymerisation of actin is required for the formation and extension of actin protrusions including lamellipodia, invadopodia and filopodia which all aid in the process of cell movement and tumour invasion (Li et al. 2008). Profilin-1 induces lamellipodia in a growth factor impendent manner by increasing the nucleotide exchange of ADP-actin for ATP-actin. In addition, genetic studies revealed that disruption of Profilin-1 gene attenuated growth and motility in single cells (Ding et al. 2006). However, below a threshold level, Profilin-1 correlated with the tumourigenic state of breast carcinoma cells (Zou et al. 2007). In contrast, in malignant melanoma cells Profilin-1 is upregulated which indicates a tumour specific role of Profilin-1 (Li et al. 2008). Its role in the progression of malignant thyroid lesions remains to be investigated. In this instance profilin-1 was down-regulated in RASV12 and upregulated in BRAF<sup>V600E</sup> and RET/PTC1 models. This implies a role of profilin in PTC tumourigenesis (Figure 14).

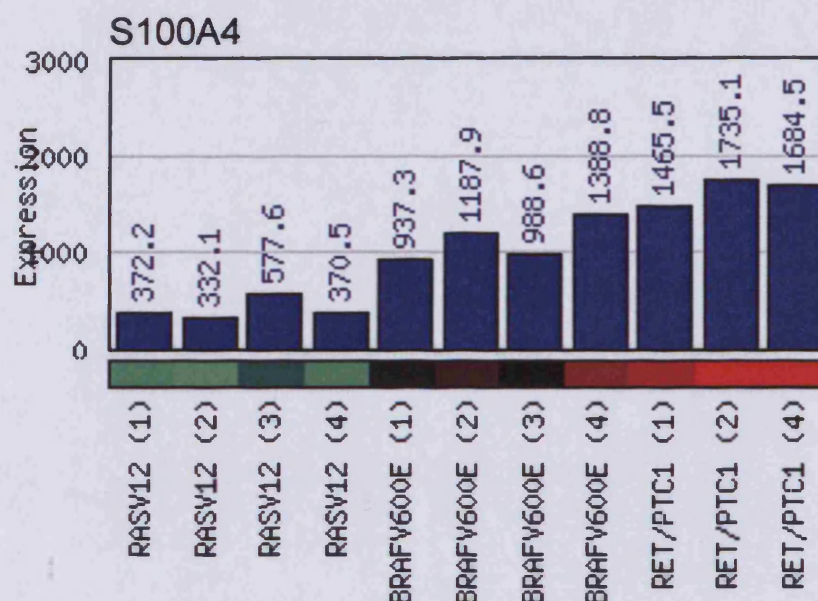




**Figure 14.** Expression analysis of Profilin-1 using MADRAS. Red highlights up-regulation, green down-regulation and black median expression

2) The S100 proteins are a family of small acidic calcium binding proteins which have a broad range of intracellular and extracellular functions including regulating protein phosphorylation and enzyme activity, calcium homeostasis, regulation of cytoskeletal components and regulation of transcription factors. They all contain 2 helix-loop-helix motifs, one of which is conserved amongst all S100 proteins and the other which conveys its specific biological function (Bjornland et al. 1999).

S100A4 has emerged as an important protein with the capacity to promote invasion and metastasis of numerous human tumours including PTC (Bjornland et al. 1999). It regulates a number of cellular processes including growth, motility, cell cycle, transcription and differentiation. Observing expression of S100A4 in MADRAS revealed basal level expression in RASV12 with increased expression in both BRAF<sup>V600E</sup> and RET/PTC1 models (Figure 15). This indicates a potential role of S100A4 in PTC.



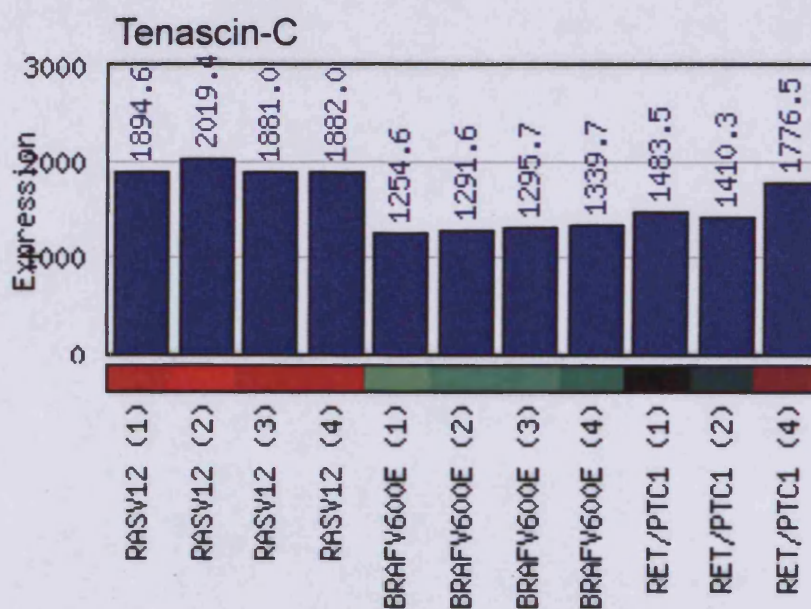
**Figure 15.** Expression analysis of S100A4 using MADRAS. Red highlights up-regulation, green down-regulation and black median expression

3) Tenascin C is an extracellular matrix glycoprotein which functions in the process of EMT. It belongs to the matricellular class of proteins which include thrombospondins and SPARC (secreted protein acidic and rich in cysteine) and can exist in soluble or insoluble forms. Tenascin C contains many repeats which share a similar homology to EGF, and fibronectin. Alternative splicing of Tenascin C gives rise to several isoforms, which may have independent roles in the extracellular matrix, although this remains to be investigated. Its expression is limited to specific developing tissues and tissues undergoing remodelling such as in wounds and tumours indicating a prevalent role of tenascin C in cell migration.

The mRNA expression of Tenascin C and fibronectin are highly upregulated in invasive malignant melanomas (Ilmonen et al. 2004). The simultaneous expression of fibronectin implies a synergistic role in carcinogenesis, a role which as yet remains



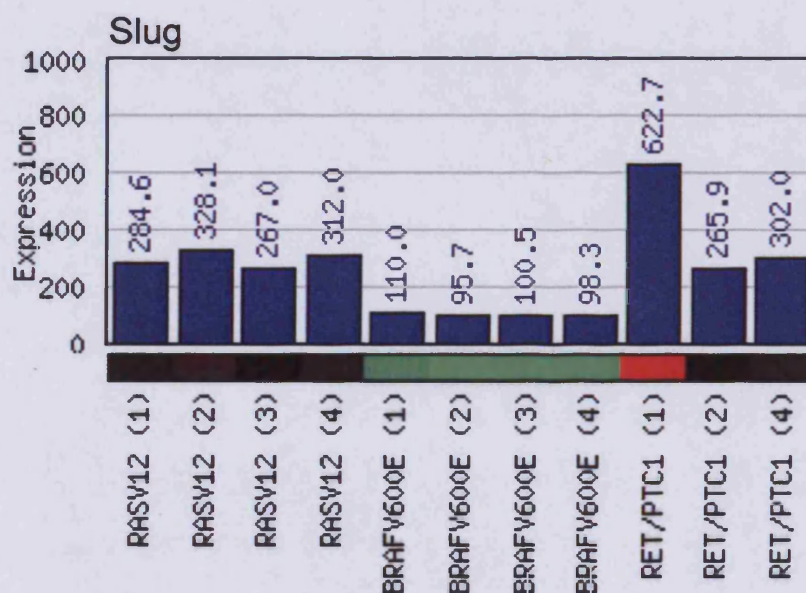
to be elucidated. Under normal conditions, fibronectin acts as an antagonist to tenascin C, with fibronectin promoting adhesion, whilst Tenascin C interferes with fibronectin-cell interactions. It is unclear if this antagonistic role by Tenascin C and fibronectin is present in thyroid lesions. In thyroid lesions, immunohistochemical analysis revealed that Tenascin C was devoid in normal and benign lesions, whilst being intensively expressed in papillary carcinomas, highlighting a potential role in malignant tumour progression (Tseleni-Balafouta et al. 2006). In my microarray profiling experiment, Tenascin-C was expressed at relatively high levels with RASV12 displaying a marginal increase in expression level (Figure 16).



**Figure 16.** Expression analysis of Tenascin-C using MADRAS. Red highlights up-regulation, green down-regulation and black median expression

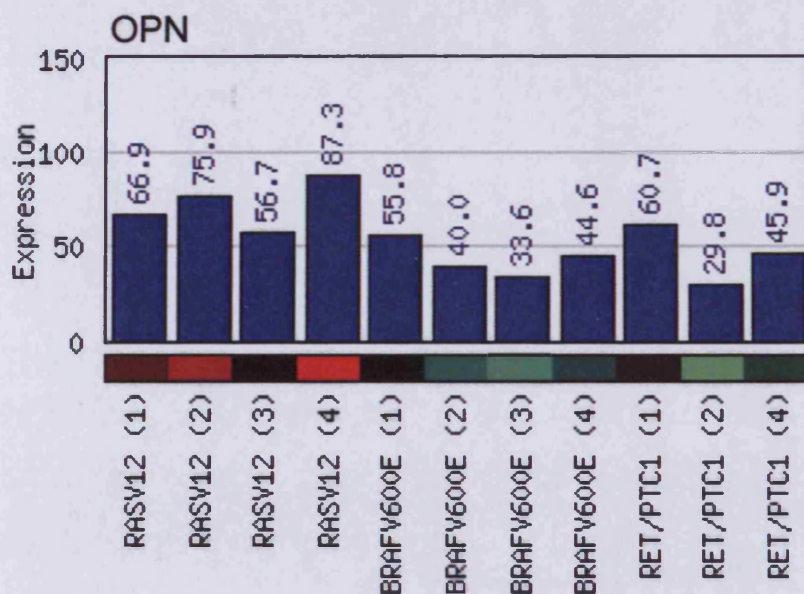


4) The Slug transcriptional repressor in recent years has become a potential target in many cancers. In normal conditions Slug expression is restricted to developmental processes during embryogenesis. In tumours it is a known transcriptional repressor of E-cadherin, which can lead to Epithelial to Mesenchymal Transition (EMT) (Hajra et al. 2002). In thyroid tumours research pertaining to Slug is limited and could be expanded upon. In my microarray profiling study the expression of Slug was detected at a median level in RASV12. RET/PTC1 also had median expression, although the RET/PTC1 (1) gene chip had double the level of expression. Expression of Slug in BRAF<sup>V600E</sup> was decreased 3-fold compared to the other models (Figure 17). In addition, the array only contained a single probe for Slug, therefore issues such as hybridisation may have compromised the probe. Further characterisation of Slug expression using alternative methods is required.



**Figure 17.** Expression analysis of Slug using MADRAS. Red highlights up-regulation, green down-regulation and black median expression

5) Osteopontin (OPN) is a secreted phosphoprotein supporting numerous cellular functions. OPN overexpression is observed in numerous malignant tumours (Wai and Kuo 2008). OPN is overexpressed in PTCs, however the mechanistic requirements of OPN in invasion requires further investigation (Guarino et al. 2005). In the microarray profiling study the expression of OPN was increased in RASV12. The expression was slightly lower the both BRAF<sup>V600E</sup> and RET/PTC1 (Figure 18).



**Figure 18.** Expression analysis of OPN using MADRAS. Red highlights up-regulation, green down-regulation and black median expression

### 3.2.8 Summary

In this section, microarrays were used to identify novel candidate genes involved with tumour invasion in BRAF<sup>V600E</sup> and RET/PTC1 models of PTC, by comparison with the RASV12 model of FA. There are many studies which utilise the microarray platform to distinguish tumour types i.e. FA from FTC. PTC can be diagnosed with 95% accuracy from FNA biopsy. My aim therefore was not to distinguish between tumour types, but rather to identify novel diagnostic markers which are implicated in tumour invasion. I conducted a microarray study using primary model systems for PTC and FA. This would provide a lower signal to noise ratio compared to primary tumours, and increased biological complexity compared to cell lines. I also aimed to provide additional biological information pertaining to the primary models.

In general, a microarray study contains a number of variables which can affect the overall outcome of experimentation and these require careful consideration when conducting profiling studies. The use of different types of platforms and gene chips may yield different expression profiles. The different use of gene chips contain within them different number of probes which therefore represent varying amounts of the human genome. The Affymetrix 133A gene chip offers 22,283 probes is compatible with its successor, the U133 PLUS 2.0 Affymetrix gene chips which were used in my experiment. These next generation gene chips offer the greatest comprehensive number of probes for human genome with an extra 6,500 genes in addition to the 22,283 probes on the 133A gene chips. This makes comparison of datasets using either of these gene chips comparable between studies. A number of studies utilise primary tumour samples or cell lines, my microarray profiling study is novel in this respect. This is the first time where transforming primary thyrocytes with respective

oncogenes creating an *in vitro* cell culture model of FA and two models of PTC. Our group is unique in our approach, and therefore is superior to cell lines and with respect to genetic profiling primary tumour samples as well. The use of cell lines offers unlimited proliferative potential with a low signal to noise ratio. The biological relevance to primary tumours is restricted as primary cells do not possess an unlimited proliferative potential, making their use as true models for primary tumours restricted, however, they remain a powerful tool in cancer research. Since primary tumour samples are taken from the said tumour, the majority of profiling studies use primary tumour tissue as the sample. However, the tumour contains a number of other cell types, in addition to the transformed tumour cells. This in effect would create 'background noise' in the experiment, perhaps masking the true profile of transformed tumour cells. The use of my group's unique primary tumour model offers an intermediate, with increased biological complexity and accuracy compared to a cell line, but decreased background noise compared to primary tumour tissue samples. In general, microarray data was analysed using 2 different approaches, the first, analysing the data by addressing the aim of the study, in this instance to identify genes which may regulate an invasive phenotype and secondly, to identify any other biological information contained within the dataset. For the identification specific genes which may regulate an invasive phenotype the *t-test* was used to compare the benign non-invasive with the invasive models BRAF<sup>V600E</sup> and RET/PTC1 which were combined for this analysis. The different kinetics in gene expression revealed by transcription expression profiling reflect the complexity of FA and PTC, represented by the RASV12 FA model, BRAF<sup>V600E</sup> PTC model and the RET/PTC1 PTC model. The second method utilised was an ANOVA test which determines differential expressed genes across all the models. This in particular displayed 5 clusters with

modulation of gene expression unique to the RASV12, BRAF<sup>V600E</sup> and the RET/PTC1 model respectively.

Since the overall aim of the microarray study was to identify novel markers mediating invasion, each oncogenic condition was analysed for their invasive behaviour using a well established *in vitro* invasion assay system. Since samples were in short supply, this particular experiment could not be conducted to its full replication with 3 separate experiments, which would normally be conducted for studies involving invasion assays. Therefore, a single experiment was conducted to represent the invasive capability of each respective model. In my hands, I found the RASV12 model displayed an invasive capability. This finding was surprising as FAs are non invasive tumours which are well known within the literature (Kondo et al. 2006). Indeed, profiling of the oncogene infected models determined a genetic profile of pathways, processes and genes which may propagate the invasive response observed in all the models.

This is the first time oncogene infected thyrocytes have been used for profiling and invasion assays. Therefore, comparisons with other studies cannot be performed. However, the normal epithelial thyrocytes which are isolated from primary tissue invariably contain multiple cell types. Although much of the unwanted cells are removed, one cannot rule out the possibility that more than one cell type is present in culture. Within the RASV12 model, our group have found a sub-culture of epithelial like cells within the primary culture. Although these cells have not been thoroughly characterised, they do display a high proliferative capacity (Bond, Wyllie et al. 1993). It is unknown if it is this sub-culture of cells which are invading the Matrigel. It is probable that the retroviral induction of RASV12 into this sub-culture conveys an invasive phenotype. Since there is no definitive detection method to determine the

invaded cell type, this can only be proposed at this stage. Introduction of mutant Ras in culture induces clonal expansion without morphological transformation, which closely represents FA *in vivo*. There has been extensive evidence which has found that introduction of mutant Ras in epithelial cells results in growth stimulation whilst in fibroblasts results in growth arrest (Gire, Marshall et al. 1999). This evidence suggests that introduction of RASV12 has different effects in different cells

Another possible explanation may be due to the culture conditions. Typically cells are cultured in culture media supplemented with 10% FCS. The presence of serum is crucial to maintain cell cycle progression, where lack of serum results in the cell cycle arrest in G<sub>0</sub> phase of the cell cycle. It is also known that serum invariably contains numerous factors including growth factors. The effect of growth factors may lead the activation and signalling of pro-invasive pathways. In the presence of serum the induction of the RASV12 oncogene may be masked by growth factors which induce the invasive phenotype which RASV12 transduced cells appear to have.

The profiling of the oncogene infected thyrocytes could not determine novel markers of invasion. This was due to the RASV12 model displaying invasive processes as determined by the microarray gene profiling experiment. Gene profiling offers an indication of processes and pathways which are implicated in invasion. In addition, in the future probable negative controls may be profiled and combined with the datasets to elucidate the aim of identifying novel markers of invasion. From the *t*-test analysis, the test is able to group both up- and down-regulated genes in the RASV12 model to compare with the combination of BRAF<sup>V600E</sup> and RET/PTC1 models. This method provided extensive gene lists, and to interrogate the lists enrichment analysis was conducted. The GeneGo Map folder and pathway map tools were utilised for the enrichment analysis. Cluster 1 which displayed probes up-regulated in



RASV12, a number of genes associated with a malignant phenotype were identified. FN is utilised as a immunohistochemical tumour marker and is found in a large percentage of carcinomas, but a small degree of adenomas also (Prasad, Pellegata et al. 2005). Detection of FN is only an indication of malignancy and cannot be considered absolute. The second gene which was identified was the pro-angiogenic factor VEGF which was also highly expressed in RASV12. VEGF has role in both physiological and pathological angiogenesis, and its expression is associated with carcinomas (de la Torre, Buley et al. 2006). Its identification in the RASV12 model implies an angiogenic response. EGFR, another gene which was expressed highly in the RASV12 model is overexpressed in anaplastic carcinomas, however, it appears posttranslational modifications convey EGFR as a differential marker for FA (Westermarck, Lundqvist et al. 1996). The identification of pathways involving Rho implies a migratory phenotype for RASV12. Rho is known to be associated with cytoskeletal reorganisation which is important for cell migration and invasion. The upregulation of this gene does in fact imply that RASV12 does not have a phenotype that represents FA. The expression of Caveolin 2 cannot determine benign or malignant phenotype in RASV12. A previous study showed Caveolin2 was not a differentiating marker of malignancy (Aldred, Huang et al. 2004). The identification of these genes indicate that RASV12 has a unique profile in that there are genes which are associated with FA, whilst a small number of genes that are associated with malignancy. It is this small proportion of genes which may contribute to the invasive phenotype of RASV12.

Genetic profiling does in part indicate why the RASV12 cells invade in the *in vitro* assay. To further expand upon this, the datasets for each oncogenic cluster were interrogated for individual probes targeted against MT1-MMP, MMP-2 and MMP-9

using MADRAS software suite. This gave a profile of MMP activity in each oncogenic condition, elucidating proteinase activity. MT1-MMP is a master regulator of MMP-2 and this was up-regulated in the RASV12 oncogenic model. Observing the expression of MMP-2, there is an indication of involvement in all the models since expression levels were comparable. Perhaps most interesting is the significant detection of MMP9 which was increased in the RASV12 oncogenic model. In immortalised keratinocytes the induction of RASV12 increases MMP-9 gene expression and promoter activity (Iyer, Pumiglia et al. 2005). Although the cellular context may be altered evidence exists which link RASV12 to MMP-9 expression, and this may explain the involvement of the protein degradation process. Although MMP gene expression was identified, the analysis does not differentiate between latent and active enzymes.

Enrichment analysis observing the map folders provided a view of the main processes and pathways occurring in each oncogenic condition. The identification of an immune response is common within the tumour microenvironment, where host tissue cells have an active participating role in tumourigenesis. The immune response links with an inflammatory response. The inflammatory response involves the influx of leukocytes and secretion of inflammatory mediators including cytokines, chemokines, growth factors and proteinases (Le Bitoux and Stamenkovic 2008).

The inflammatory process was identified in all the oncogenic models; however, a normal thyroid negative control is required to observe if the involvement of inflammation and immune system response is involved in invasion. Tumours whether malignant or benign, are subject to the host immune-surveillance. This is where the host immune system responds to transformed cells by inducing an immune response, as they are deemed a foreign entity. Although this may occur tumours also



possess mechanisms to evade the host immune response. Deficiencies in key immunological molecules may convey the identification of the immune system response and inflammation in the analysis.

I also sought to determine the expression levels of targets I had identified within the literature. Based on my extensive literature analysis and expression profiling experiment I aimed to select a couple of candidates to take forward for further experimentation.

From the literature analysis I had selected a number of genes. These included the transcription factor Slug, OPN, profilin-1, Tenascin-C and S100A4. All of these genes have a role in tumourigenesis, or are linked to invasive processes. The S100A4 protein is a small calcium binding protein that is associated with metastatic tumours and appears to be a molecular marker for clinical prognosis. The expression of S100A4 is therefore a good indication of an invasive phenotype. Since it is associated with thyroid cancer, it is not only a probable candidate for genetic manipulation, but also an indicator of the gene profile of the oncogenic models. S100A4 is highly expressed in the PTC models, a result which is consistent with the models and the literature. There has been a vast amount of work conducted with the S100A4 in thyroid cancer, and in my opinion investigating this particular gene would not provide much novel work for this thesis.

Profilin-1 was highly expressed in my PTC models, however, in breast cancer cells profilin-1 acts as a suppressor of tumourigenesis (Janke, Schluter et al. 2000). Therefore this gene is not a strong choice for genetic manipulation. Tenascin-C expression in the *in vitro* models revealed similar levels. In melanoma, tenascin-C is associated with tumour progression (Fukunaga-Kalabis, Martinez et al.). Without

significant differential expression tenascin-C remains a weak candidate for further investigation.

However, OPN is known to be overexpressed in PTC (Guarino, Faviana et al. 2005) and details regarding the invasive mechanisms regulated by OPN remain elusive. The genetic profile of OPN revealed a low level of detection in the models. This implies the PTC models do not have an 'exact' genetic profile to the *in vivo* situation. The expression profile of OPN was also variable with differing degrees of detection between each chip. There was only a single probe for OPN on the gene chip and this may explain the variable result seen. From the literature genetic manipulation of OPN has not been conducted before, therefore the scope for elucidating OPN in thyroid tumour invasion is great.

The Slug transcription factor is well established in tumours and has been implicated in thyroid tumourigenesis (Nieto 2002). Genetic manipulation of this gene has not been conducted before, and observing expression levels in the models revealed median expression in both RASV12 and RET/PTC1. The levels were similar between the models, and all had displayed an invasive capability, therefore this gene is a strong contender for genetic manipulation.

Taken together, genes which were identified in the extensive literature search were selected as candidate markers for thyroid invasion. The OPN and Slug genes were selected as candidates for genetic manipulation based on their implicated role in thyroid tumourigenesis, and their scope for future work. Although the expression values obtained in MADRAS are an indicator of a possible role, this cannot be taken as absolute without a control displaying a non-invasive phenotype, or re-confirming the expression using real-time quantitative PCR.

# **PART II**

## **4.0 Background**

## 4.1 The SNAIL family of transcriptional repressors

In vertebrates, the Snail family of zinc finger transcription factors consists of 3 genes, *snai1* (snail), *snai2* (slug) and *snai3* (smuc). The Snai1 zinc finger transcription factor was first identified in *Drosophila melanogaster*, where it is required for mesoderm development (Boulay, Dennefeld et al. 1987; Alberga, Boulay et al. 1991). The family all have a conserved carboxy-terminal containing 4-6 zinc fingers to function as DNA binding motifs, with a diverse amino-terminal domain. The consensus binding site for snail-related genes is identical to the E-box binding site, a core binding site for basic-helix-loop-helix binding transcription factors (Nieto 2002). The binding of the E-box by Snail family members relies upon the zinc finger domains and the SNAG domain contained within the amino terminal (Nakayama, Scott et al. 1998). The central domain is a divergent serine/proline rich region, where the Slug protein contains a so-called Slug domain within this region. Within the Snai1 protein, the central portion contains a regulatory nuclear export signal (NES) domain and a destruction box domain (Peinado, Olmeda et al. 2007).

The process of EMT is important for early development during gastrulation and neural crest cell migration which require motile cells that differentiate into specialised structures. In addition in the context of tumour progression in general, EMT is initiated through the disruption of cell-cell contacts which include adherent, tight and desmosomal junctions that link epithelial cells to one another. This is followed by cytoskeletal reorganisation and degradation of basement membranes and stroma together with an enhanced growth rate (Kurrey, K et al. 2005). The repression of the adherent Epithelial-cadherin (E-cadherin) protein is important for EMT. It functions as a calcium dependant adhesion protein and facilitates cell-cell communication and is

regarded as a tumour suppressor. The cytoplasmic domain of E-cadherin interacts with  $\alpha$  and  $\beta$  and p120ctn catenins and this complex interacts with the actin cytoskeleton. The extracellular domain binds to adjacent adherent junctions on neighbouring cells.  $\beta$ -catenin is associated with cell-cell adhesion but is also a well established mediator of the Wnt signalling pathway. Disruption of cell-cell adhesion includes the downregulation of E-cadherin, and this is considered as a progressive step in conferring invasive potential of carcinomas. E-cadherin is repressed by the snail family of transcription factors, which can interact with E-box elements on the E-cadherin promoter (Cano, Perez-Moreno et al. 2000). Indeed, Snai1 is expressed in the developing skin of mice, where the skin cells lose E-cadherin expression during the formation of hair follicle buds (Jamora, Lee et al. 2005). Also, Snail mutant mice die at gastrulation which highlights the critical role Snail in EMT (Sefton, Sanchez et al. 1998). Snai1 silencing has also been shown to inhibit tumour growth (Olmeda, Jorda et al. 2007). The event of Snai1 overexpression in Madin Darby canine kidney (MDCK) epithelial cells leads to an invasive phenotype (Cano, Perez-Moreno et al. 2000). Snai1 expression has been highlighted in an increasing number of human carcinoma and melanoma cell lines (Peinado, Portillo et al. 2004), and is expressed at the invasive front of epidermoid carcinomas (Cano, Perez-Moreno et al. 2000) and has been correlated to lymph node invasion of ductal breast carcinomas and hepatocarcinomas (Blanco, Moreno-Bueno et al. 2002; Sugimachi, Tanaka et al. 2003). The high expression of the Snail transcription factors implies a role in thyroid carcinogenesis. Increased expression was detected in thyroid papillary carcinoma cell lines, with lack of expression in normal thyroid tissue or normal tissue derived cell lines. This particular study demonstrated this was due to the downregulation of E-cadherin (Hardy et al., 2007).

The second member of the Snail transcription factors, Slug has an overlapping role with Snail but also has been linked with additional EMT processes. Slug has been documented to be a repressor of E-cadherin (Hajra, Chen et al. 2002; Uchikado, Natsugoe et al. 2005), but also has been shown to convey other cellular processes which are important for EMT. Indeed, cell survival is conveyed by Slug during epithelial morphogenesis (Leroy and Mostov 2007) and has been shown to function as a survival factor in myeloid progenitor cells, where Slug repressed the p53-induced apoptotic gene puma, conferring cell survival (Wu, Heinrichs et al. 2005). In addition to E-cadherin repression, Slug in numerous pathological contexts has been shown to be responsible for desmosome dissociation where E-cadherin was not repressed (Savagner, Yamada et al. 1997). In epithelial ovarian cancers, in addition to the downregulation of E-cadherin, Snail transcription factors also repress other junctional proteins including Claudin, Occludin and ZO-1 (Bolos, Peinado et al. 2003; Ikenouchi, Matsuda et al. 2003; Ohkubo and Ozawa 2004) and regulation of other epithelial markers including cytokeratin 18 and Mucin-1 have been demonstrated (Guaita, Puig et al. 2002; Tripathi, Misra et al. 2005). Cytokeratins are intermediate filament proteins present on epithelial cells and their downregulation indicates biochemical and morphological changes often seen in carcinomas. Slug has also been shown to decrease the expression of integrins  $\alpha 3$ ,  $\beta 1$  and  $\beta 4$  in cultured epidermal keratinocytes (Turner, Broad et al. 2006). In an ovarian epithelial cancer cell line both expression of Slug and E-cadherin was detected, indicating an alternative role of Slug in tumourigenesis (Kurrey, K et al. 2005). It seems then, that Slug expression is dependent on the pathology of the tumour and its associated EMT process. In the context of bone formation or osteoblast maturation, Slug positively correlated with osteoblast markers including Runx2, OPN, collagen type I

and Wnt signalling mediators and the knockdown of Slug weakened the Wnt signalling cascade, as well as reducing the other osteoblastic markers (Lambertini, Lisignoli et al. 2009). This highlights the broad range of functions in which Slug mediates and partakes.

From embryogenesis, Snai1 appears to be regulated by the NFκB and the FGF receptor. Transcription of Snai1 requires ERK signalling, indicating the MAPK pathway as the transcription can be influenced by MEK inhibitors (Barbera, Puig et al. 2004). The PI3K pathway has also been shown to regulate Snail expression, although it may not have such an integral role (Barbera, Puig et al. 2004). The repression of E-cadherin induces translocation of β-catenin to the nucleus where it binds to the Lcf transcription factor to aid in EMT (Conacci-Sorrell, Simcha et al. 2003; Medici, Hay et al. 2008). In addition, TGFβ1 and TGFβ2 have both been shown to promote Snail transcription, highlighting the upstream elements of Snai1 and Slug. Also, TGFβ3 was shown to induce Lcf, which highlights the crossover between the TGFβ and Wnt signalling pathways (Medici, Hay et al. 2008). Correlating with these findings TGFβ1 was shown to induce Snai1 expression in epithelial cell lines (Peinado, Quintanilla et al. 2003). This also applies to Slug, where treatment of TGFβ1 induced Slug expression in lens and epithelial cells *in vitro*, and Slug expression was reduced with MEK inhibitors (Choi, Park et al. 2007). This indicates that both Snail transcriptional repressors function on the same pathway

Matrix metalloproteases (MMPs) have important roles in the degradation of the ECM, aiding invasion and tumour progression. The expression of MMP-2, which degrades type IV collagen is prevalent in Slug-transfected cells (Shih, Tsai et al. 2005). In malignant mesothelioma, Snail protein expression was significantly associated with

the expression of MT1-MMP (Sivertsen, Hadar et al. 2006). Snai1 has also been reported to increase the expression of MMP-2 (Yokoyama, Kamata et al. 2003).

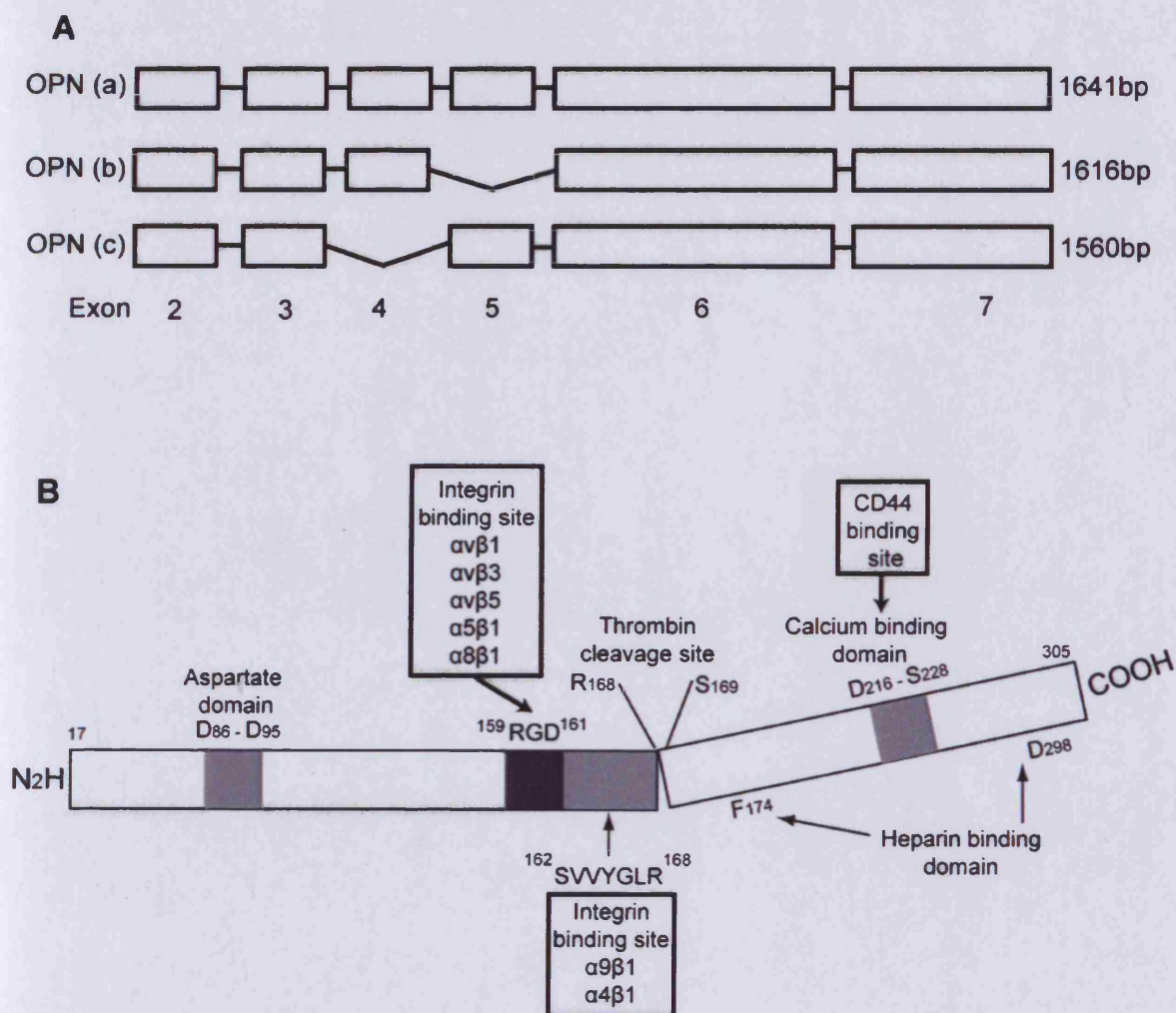
## **4.2 Osteopontin (OPN)**

Osteopontin (OPN) or secreted phosphoprotein 1 is localised on chromosome 4 and belongs to the family of small integrin binding ligand N-linked glycoproteins, which also includes in addition to OPN, bone sialoprotein, dentin matrix protein 1, dentin sialoprotein and matrix extracellular phosphoglycoprotein (Wai and Kuo 2004). OPN was first characterized as a phosphoprotein secreted by transformed epithelial cells (Senger, Wirth et al. 1979). OPN is expressed in multiple species and in multiple cell lines including osteoclasts, osteoblasts, epithelial cells, nerve cells, vascular smooth muscle cells and endothelial cells (Wai and Kuo 2004). Its expression has also been highlighted in activated immune cells including T-cells, macrophages and natural killer (NK) cells (Wai and Kuo 2004). Human OPN maps to chromosome 4q13 and contains 7 exons, with the secreted form between 44 kDa – 75 kDa in size. This is thought to be due to posttranslational modifications by O-glycosylation, sulfation, phosphorylation and enzymatic cleavage, allowing secreted OPN to support its numerous cellular functions. There is evidence for the alternative splicing of the OPN gene, where 3 cDNAs have been identified. These include the full length OPN (a) gene, OPN (b) lacking exon 5 and OPN (c) lacking exon 4. The functional significance of these deletions is currently unknown (Figure 19).



#### *4.2.1 Structure of OPN*

A protease sensitive site within OPN separates its binding domains with integrins and CD44. Its thrombin cleavage motif contains a conserved sequence known as RSK. Thrombin cleaves OPN at this motif releasing two receptor binding domains. The secreted OPN protein is incorporated in the matrix in mineralised connective tissue but also is detected in biological fluids including bile, urine sweat and semen (Ramaiah and Rittling 2008); (Senger, Perruzzi et al. 1988); (Senger, Perruzzi et al. 1989); (Bautista, Denstedt et al. 1996).



**Figure 19.** A. Genomic structure highlighting OPN spliced isoforms. B Schematic representation of the domain structure of OPN with binding ligands (black boxes). (Rangaswami, Bulbule et al. 2006); (Denhardt, Noda et al. 2001); (Ramaiah and Rittling 2008).

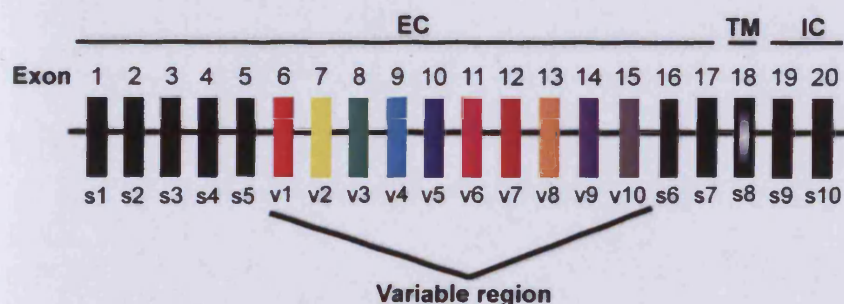
OPN has been characterised as a secreted glycoprotein, however an intracellular form of OPN exists. This was detailed in plasmacytoid dendritic cells (Shinohara, Kim et al. 2008), fetal fibroblasts (Zohar, Suzuki et al. 2000) and peritoneal macrophages (Zhu, Suzuki et al. 2004). This isoform is not subjected to post-transcriptional modifications and is synthesised independently of secreted OPN (Zohar, Suzuki et al. 2000). Intracellular OPN can interact with CD44. This interaction indicates a probable role in increased cell migration, as disruption of intracellular OPN or CD44 causes a reduction in chemotaxis (Zhu, Suzuki et al. 2004).

The main integrin binding sites are situated on the N-terminal cleaved fragment of OPN. The RGD domain is a binding site for numerous integrins  $\alpha v/\beta 3$ ,  $\alpha v/\beta 1$ ,  $\alpha 5/\beta 1$ ,  $\alpha 8/\beta 1$  and  $\alpha v/\beta 5$ , whilst the SVVYGLR domain can bind  $\alpha 9/\beta 1$  and  $\alpha 4/\beta 1$  integrins. The SVVYGLR domain is cryptic and is only accessible when OPN is cleaved by thrombin. It is also a substrate for MMP-3 and MMP-7 cleavage, providing an additional level of signalling capability (Ramaiah and Rittling 2008). It is currently unreported which isoforms of OPN are secreted.

## 4.2.2 OPN receptors

### 4.2.2.1 CD44

The CD44 family of molecules are single chain glycoproteins encoded by a single copy gene which is located on the short arm of chromosome 11p13 and is composed of 20 exons (Jothy 2003). Its function include cell adhesion affecting aggregation and migration, lymph node homing, myelopoiesis and lymphopoiesis, angiogenesis and release of cytokines (Sneath and Mangham 1998). The protein size vary between 80 and 200 kDa, which is due to post-transcriptional modifications including glycosylation and alternative splicing (Marhaba and Zoller 2004). The formation of exons 1-5 and 16-20 forms the ubiquitously expressed standard isoform of CD44 (CD44s) (Goodison, Urquidi et al. 1999). Indeed the alternative splicing allows the insertion of variable exons between exons 5 and 16 located within the extracellular, membrane proximal region of CD44 (Marhaba and Zoller 2004) and are designated CD44v giving further diversity to the functional regulation of CD44. The human CD44 mRNA, unlike mouse does not contain exon 6 due to the presence of a stop codon. post-transcriptional splicing may result in up to 10 isoforms of CD44 (Figure 20).



**Figure 20.** Exon map of CD44 displaying spliced exons and conserved exons. Spliced variants are named according to the corresponding v exon splicing. Exons v6, v7 and v 10 are prominent in malignancy.

The principle ligand for CD44 is hyaluronic acid (HA) which is ubiquitous component of the ECM. There are a total of 3 binding domains for HA on CD44 at the exon 2 link domain and 2 domains encoded by exon 5 (Goodison, Urquidi et al. 1999). Mutation studies have shown that the binding sites consist of basic amino acid clusters with arginine residues (Bajorath, Greenfield et al. 1998). In some instances in response to an external stimulus CD44 can bind other molecules including galectin-8, collagen, fibronectin, laminin, fibrinogen, chondroitin sulphate, mucosal vascular addressin, serglycin/gp600 and OPN (Naor, Wallach-Dayana et al. 2008). The C-terminal fragment of OPN interacts with CD44, with an interaction also shown to occur at the N-terminal fragment independent of the RGD domain of OPN (O'Regan 2003).

The expression of variant isoforms of CD44 have been implicated in tumour progression (Jothy 2003). However, this seems to be tumour specific, for example, in prostate carcinoma the down-regulation of CD44v6 correlates with metastasis (Ekici, Ayhan et al. 2002), and invasion and migration is inhibited through the overexpression of CD44s and knockdown of CD44v7-10 (Omara-Opyene, Qiu et al. 2004). Although CD44v are membrane bound, there is evidence that they exist in soluble forms in the serum of gastrointestinal carcinomas, cervical carcinomas and malignant lymphomas (Guo, Liu et al. 1994; Kainz, Tempfer et al. 1995; Ristamaki, Joensuu et al. 1997). Indeed serum levels correlate with tumour burden in colon cancer and malignant lymphoma making soluble CD44 a prognostic marker in certain tumours. Soluble CD44 is formed by proteolytic cleavage or CD44 shedding. Shedding of CD44 occurs through proteolytic cleavage of its ectodomain, and release of the extracellular domain (Okamoto, Kawano et al. 1999). CD44 shedding has been highlighted in numerous tumours *in vivo* (Okamoto, Kawano et al. 1999). The cleavage of CD44 was detected in 58% of gliomas, 67% breast carcinomas,

45% of non-small cell lung carcinomas, 90% of colon carcinomas and 25% of ovarian carcinomas (Okamoto, Tsuiki et al. 2002). Proteolytic cleavage of CD44 was also found to occur in highly aggressive melanoma cell lines and may confer an invasive phenotype (Goebeler, Kaufmann et al. 1996). In cervical cancer cells, CD44 is known to interact with the EGF Receptors, erbB2 and erbB3 enhancing their signalling during cervical carcinoma progression by constitutive activation (Wobus, Kuns et al. 2001).

The shedding of CD44 is mediated by membrane bound matrix metalloproteinases. Furthermore, ADAM metalloproteinases have also been shown to induce CD44 cleavage (Nakamura, Suenaga et al. 2004). This is triggered in response to numerous stimuli including protein kinase C signalling activation, calcium influx, Ras and Rac GTPase activity (Okamoto, Kawano et al. 1999) and EGF (Palyi-Krek, Barok et al. 2008). With respect to Ras involvement, it was shown that H-Ras transformation in NIH3T3 fibroblasts results in increased CD44 cleavage increasing cell migration and invasion of the ECM (Kawano, Okamoto et al. 2000). In breast cancer, the shedding of CD44v6 is associated with increased tumour size, lymph node metastasis and resistance to chemotherapy. The cytoplasmic tail of CD44 can interact with many cytoskeletal proteins, namely the band 4.1 family of proteins which are ezrin-radixin-moesin (ERM) family and merlin. The ERM proteins are integral to the regulation of cell migration, cell shape and protein resorting at the plasma membrane. This complex rests in an inactive state, and become activated to link the transmembrane proteins with the actin cytoskeleton (Marhaba and Zoller 2004).

Proteolytic cleavage occurs in the extracellular domain which produces an extracellular CD44 fragment, it is followed by cleavage of the CD44 intracellular

domain, mediated by  $\gamma$ -secretase, a membrane bound protease complex (Steiner, Fluhner et al. 2008). This leads to translocation to the nucleus which interacts with the p300/cAMP responsive element to initiate the transcription of a number of genes including CD44 itself, thus providing a positive feedback loop (Nagano and Saya 2004). The adenomatous polyposis coli (APC) gene mutation in colorectal cancer mutation is well known. This gene is integral to the WNT signalling cascade interacting with  $\beta$ -catenin and indirectly interacts with cadherin molecules (Wielenga, Smits et al. 1999). In intestinal epithelial cells, mutation leads to constitutive transcription of the Tcf-4. The activation of Tcf-4 signalling which occurs with APC mutations is associated with CD44 overexpression. Targeted blockade of Tcf-4 abrogates CD44 expression (Wielenga, Smits et al. 1999). This interaction implies Tcf-4 either directly or indirectly regulating CD44 expression in colorectal cancer (Wielenga, Smits et al. 1999). Inhibition of CD44 in mammary cancer cells induces apoptosis highlighting a role of CD44 in the process of programmed cell death, and overexpression may lead to a counteraction of this process (Yu, Toole et al. 1997).

The interaction of OPN and CD44 has been shown to confer increased cell survival (Lin, Huang et al. 2000; Lee, Wang et al. 2007) . In H-RasV12 transformed cells, both OPN and CD44 are upregulated inducing an autocrine feedback loop, and this autocrine loop is directly related to H-RasV12 cell invasion in the system (Teramoto, Castellone et al. 2005). CD44v6 and 7 are reported to interact with OPN (Katagiri, Sleeman et al. 1999; Lee, Wang et al. 2007). This interaction was shown to correlate with the clinical outcome of laryngeal dysplasia (Staibano, Merolla et al. 2007) and confer cell survival in gastrointestinal cancer cells (Lee, Wang et al. 2007). This type of interaction has also been established in other carcinomas (Ponta, Sherman et al. 2003). The OPN-CD44 is not strictly exclusive, since in rat pancreatic carcinoma  $\beta$ 1

integrins cooperate with CD44v to promote OPN-dependent cell motility and chemotaxis (Katagiri, Sleeman et al. 1999).

#### *4.2.2.2 $\alpha v \beta$ Integrins*

OPN can bind to  $\alpha v \beta 3$ ,  $\alpha v \beta 1$ ,  $\alpha v \beta 5$  integrin through the RGD domain. The  $\beta 1$  integrins have been shown to cooperate with CD44v mediated OPN to increase cell motility and chemotaxis (Katagiri, Sleeman et al. 1999). Integrin  $\beta 3$  is consistently detected in breast cancer, and mediates OPN migration and invasion of tumour cells. VEGF, a pro-angiogenic factor can induce OPN and  $\alpha v \beta 3$  expression in microvascular endothelial cells to promote angiogenesis (Wai and Kuo 2004). OPN- $\beta 3$  integrin interaction mediates cellular adhesion of a number of cell types (Wai and Kuo 2008). This interaction also detected in breast cancer metastasis, and mediates invasion and cell migration of tumour cells (Furger, Allan et al. 2003).

The distribution of integrin  $\alpha v \beta 3$  is polarised towards the cell front in migrating fibroblasts (Woods, White et al. 2004). In epithelial cells integrin  $\alpha v \beta 3$  is persistently expressed and appears to be the result of Rac activation (Danen, van Rheenen et al. 2005). The overexpression of integrin  $\alpha v \beta 3$  increases Rac activity and lamellipodium formation, increasing actin stress fibre formation. Engagement of integrin  $\alpha v \beta 3$  results in integrin clustering and the formation of Rac-dependent focal complexes which are important for cell migration (Kiosses, Shattil et al. 2001).



### 4.2.3 Regulation of OPN expression

#### 4.2.3.1 *Ras*

Direct transcriptional regulation of OPN is mediated by *Ras*, a GTPase which primarily activates the MAPK pathway. An increase in OPN transcription is observed in *Ras*-transformed cells, and knockdown of OPN in these cells resulted in a reduced ability to form tumours and metastasise (Denhardt, Mistretta et al. 2003). Conserved among species, the (murine) OPN promoter contains a *Ras*-activated enhancer domain and confers increased transcriptional activity of the OPN gene in *Ras*-transformed 3T3 cells (Zhu, Denhardt et al. 2005). In addition, the haematological malignancy chronic myeloid leukaemia (CML) is caused by the constitutive active fusion gene Bcr-Abl. In response to high levels of Bcr-Abl, a fusion gene product in self-renewing haematopoietic stem cells OPN is upregulated. The signalling pathways mediating this response require *Ras*, atypical PKC, MEK and Raf-1 activation in CML (Hickey, England et al. 2005).

#### 4.2.3.2 *Rous sarcoma oncogene (Src)*

The proto-oncogene, *Src* is a non receptor tyrosine kinase which has roles in proliferation, adhesion and survival (El-Tanani, Campbell et al. 2006). *Src* actively regulates OPN in tumour and non-tumour cells, where v-*Src* and c-*Src* induce OPN expression in transformed NIH3T3 mouse fibroblasts (Tezuka, Denhardt et al. 1996). This is almost certainly to be due to a v-*Src* response element in the murine OPN promoter (Blobel and Hanafusa 1991). The v-*Src* tyrosine kinase induces several

genes including collagens, TGF $\beta$ , c-fos and MMP-9, and may aid the process of tumour invasion and metastasis (Wai and Kuo 2008).

#### 4.2.3.3 *Tcf-4*

The Wnt signalling pathway has key roles in development, homeostasis and tumour progression (Hatsell, Rowlands et al. 2003). Detailed description of the pathway has been provided earlier. Briefly, the pathway is activated by Wnt proteins which bind and activate the membrane bound Frizzled receptors, which mediate the phosphorylation of dishevelled (Dsh).  $\beta$ -catenin is a key regulator of the Wnt signalling pathway. When phosphorylated by the APC/axin/GSK3 $\beta$  complex, it is primed for destruction. Unphosphorylated  $\beta$ -catenin translocates to the nucleus to bind the Lcf/Tcf family of transcription factors to initiate activation of a host of genes including c-myc, cyclinD1 and OPN. In the absence of  $\beta$ -catenin Tcf-4 binds to A/TATCAAAG sequences in the OPN promoter inhibiting OPN protein expression. In the presence of  $\beta$ -catenin the effect is opposite with OPN promoter activation and protein expression (El-Tanani, Barraclough et al. 2001). Deregulation of the Wnt pathway has been identified in colorectal cancer, which results in increased active  $\beta$ -catenin levels in both the cytoplasm and nucleus, and OPN expression significantly correlates with this expression highlighting that the Wnt pathway is largely responsible for OPN activation and expression (Rohde, Rimkus et al. 2007).

#### 4.2.3.4 *AP-1*

The AP-1 transcription factor is a dimeric complex which mediates transformation in a number of cellular contexts including bone, skin and in lung tumours. The OPN

promoter also contains a binding site for AP-1 in its promoter region (Renault, Jalvy et al. 2003). AP-1 can induce expression of OPN by binding to its promoter (El-Tanani, Platt-Higgins et al. 2004). A positive feedback mechanism exists where OPN can stimulate AP-1 transactivation and DNA binding mediated by  $\alpha\beta 3$  integrin and ERK signalling (Das, Mahabeleshwar et al. 2004). The binding of  $\alpha\beta 3$  integrin activates integrin linked kinase (ILK) which leads to the binding of AP-1 to the proteinases MMP-2 and uPA promoters and an increase in promoter activity (Mi, Guo et al. 2006).

#### **4.3.4 OPN and proteinase activity**

OPN has important roles in many physiological cellular functions, including mineralisation, blood vessel formation and inflammation. Cumulative evidence has highlighted the role of OPN in tumourigenesis and metastasis. The remodelling of the cytoskeleton represents a major event during tumour invasion. OPN mediates the regulation of proteinases to regulate the degradation of the extracellular matrix. The urokinase plasminogen activator (uPA) has a role in tumour progression through regulation of cell motility and extracellular matrix degradation. OPN has been shown to induce uPA secretion and subsequent activation of plasmin, thereby regulating tumour cell migration, motility and extracellular matrix degradation in breast cancer cells. The induction of uPA occurs via c-Src, PI3K and MAPK signalling (Das, Mahabeleshwar et al. 2003). In addition, OPN mediates the expression of proMMP-2 during tumour invasion in an NF $\kappa$ B-dependent manner in melanoma cells (Philip, Bulbule et al. 2001). In chrondrasarcoma cells, OPN mediates the expression of MMP-9 through a number of signalling mechanisms including  $\alpha\beta 3$  integrin, FAK and

MAPK signalling, all of which indicate a role in mediating the expression of MMP-9 and cell migration (Chen, Wei et al. 2009). In prostate cancer, OPN regulates MMP-2 and COX-2 expression to mediate invasion and motility *in vivo* (Jain, Chakraborty et al. 2006). COX is an integral membrane bifunctional enzyme which metabolises arachidonic acids to active eicosanoids including PGE<sub>2</sub> (Jain, Chakraborty et al. 2006). Increased COX-2 expression has been detected in a number of tumour pathologies including colorectal, gastric, pancreatic and breast and others (Wallace 2002) and generation of PGE<sub>2</sub> contributes to the promotion of angiogenesis, invasive potential, motility and survival (Surh and Kundu 2005).

#### **4.2.5 The Hepatocyte Growth Factor (HGF)-cMet system**

One of the mechanisms that OPN confers a migratory phenotype in tumour cells is through the HGF-cMet system. The cMet tyrosine kinase receptor is predominantly expressed on epithelial and endothelial cells, its ligand, hepatocyte growth factor (HGF) is produced by motile mesenchymal cells. The interaction leads to a number of physiological functions including development of epithelial organs, tissue regeneration, angiogenesis and wound healing (Birchmeier and Gherardi 1998). In an *in vitro* setting, HGF is a potent stimulus for cell growth, motility and morphogenesis. Deregulation of this signalling system is implicated in numerous human cancers (Ariztia, Subbarao et al. 2003). Mutation of cMet may cause ligand independent receptor activation and cMet expression (Ariztia, Subbarao et al. 2003). OPN is prominent in HGF-dependent tumourigenesis, and inhibition of OPN expression *in vitro* and *in vivo* attenuates the HGF-induced (Medico, Gentile et al. 2001; Ariztia, Subbarao et al. 2003). This is an indication that OPN is a necessity for

HGF-induced tumour invasion. CD44 variant isoforms are able to induce the amplification of HGF induced invasion. OPN can interact with CD44v together with  $\beta 1$  integrins to induce receptor pair clustering (van der Voort, Taher et al. 1999). This mechanism is thought to occur in a number of possible ways, firstly, both CD44v and  $\beta 1$  integrins binding OPN in a cooperative manner; secondly, signalling via CD44v mediating the OPN- $\beta 1$  integrins; and finally, by interacting with  $\beta 1$  integrins and stabilising the CD44v interaction (Ariztia, Subbarao et al. 2003). In addition, the CD44v6 and CD44v7 have been shown to bind to OPN to enhance the migratory phenotype of rat fibrosarcoma cells (Katagiri, Sleeman et al. 1999).

#### **4.2.6 OPN in thyroid cancer**

In thyroid cancer, OPN is overexpressed and enhances papillary carcinoma invasiveness in RET/PTC transformed rat thyroid cells. In thyroid tumours OPN expression correlates with aggressive clinicopathological features of papillary carcinoma, presence of lymph node metastases and tumour size. OPN-integrin binding directly mediates migration and invasion of tumour cells (Angelucci et al., 2002), and is known to enhance thyroid papillary carcinoma invasiveness (Guarino, Faviana et al. 2005). Moreover, in RET/PTC transformed cells, an autocrine loop sustaining proliferation and invasiveness initiated by OPN-CD44 signalling was discovered (Castellone, Celetti et al. 2004).

### **4.3 Tumour Invasion and Metastasis**

Growth of primary tumour is often followed by the metastasis of tumour cells and formation of secondary tumours to distant sites of the body. For this to occur, a number of events need to occur for tumour cells to invade and disseminate neighbouring tissues, and for the tumour cells persist and progress in secondary sites. The progression of cancers may be considered as a collection of cellular processes converging together to induce invasion, a prominent feature of cancers and their metastases. The metastatic process consists of a series of events which must occur for a cancer to progress (Chambers, Groom et al. 2002). As a primary tumour grows, angiogenesis needs to occur so the tumour can support its metabolic needs and to provide a metastatic route. Cells in the primary tumour invade the surrounding tissue stroma and migrate towards blood/lymphatic systems. Intravasation into blood vessels allows the tumour cells to disseminate to distant parts of the body using the host circulatory system. Here the cancerous cells may persist in the vessels or extravasate and invade other organs.

#### **4.3.1 Epithelial to mesenchymal transition (EMT)**

In normal conditions, cell migration and motility are essential for tissue development and homeostasis. Tumour cells acquire these processes to further their progression to aid invasion, metastasis and survival. The epithelial-mesenchymal transition (EMT) naturally features during embryogenesis (Hugo, Ackland et al. 2007) wound healing, (Desmouliere 1995) placental formation (Vicovac and Aplin 1996) and production of fibroblasts during inflammation (Iwano, Plieth et al. 2002). It essentially brings a change in cellular morphology to allow cells to have a migratory phenotype. Alteration of an epithelial cell to have mesenchymal properties requires morphological, architecture, adhesion and migratory changes. Normal epithelial cells form a tight monolayer and their movement is restricted and provide tissues with mechanical rigidity. The presence of cell-cell junctions and adhesive molecules holds the cells tightly together and inhibits the movement of individual cells. In contrast, mesenchymal cells do not possess the epithelial junctions and their adhesive properties are less strong allowing for greater motile capacity. They form structures which are irregular in shape and not uniform in structure or density. The morphology of mesenchymal cells is irregular as they possess leading edge polarity. The difference between epithelial and mesenchymal cells defines an essential dynamic process for tumour progression, and EMT is a crucial event for this to occur. The transition involves a change in the adhesive qualities of the cell, remodelling of the cytoskeleton, and proteolytic degradation of the extracellular matrix.

### **4.3.2 Cell adhesion and cytoskeletal rearrangement**

Cell migration has an integral role in a diverse range of biological functions including development, immune regulation, and tumourigenesis (Dikeman, Rivera Rosado et al. 2008). In the context of tumour progression, tumour cells need to detach from the primary tumour mass, inhibit cytokinesis, initiate protrusion formation, cell migration and remodelling the surrounding ECM through by physical and chemical means.

#### **4.3.2.1 Cell Adhesion**

##### *A) Integrins*

For directional movement of cells the spatial and temporal coordination of membrane protrusions and retraction with cell adhesion are required. Indeed, cell migration is coordinated by numerous proteins which regulate cell-matrix and cell-cell interactions, formulating structural linkages or adhesions with cytoskeletal proteins to become organised protein complexes. Central to interactions are the family of integrins which formulate the bridge between the ECM and the actin cytoskeleton of the cell. Integrins are multifunctional heterodimeric adhesion receptors consisting of non-covalently bound  $\alpha$ - and  $\beta$ - subunits. Currently, there are 18 different  $\alpha$ - subunits and 8  $\beta$ - subunits forming a total of 24 distinct heterodimers (Al-Jamal and Harrison 2008). The general structure of an integrin subunit consists of a large ectodomain, a transmembrane domain and a short cytoplasmic tail. The globular head is devoid of kinase activity and binds ECM components.

The activation of integrins is important for cellular processes involving anchorage such as cell polarity of migrating cells and ECM assembly (Ginsberg, Partridge et al.

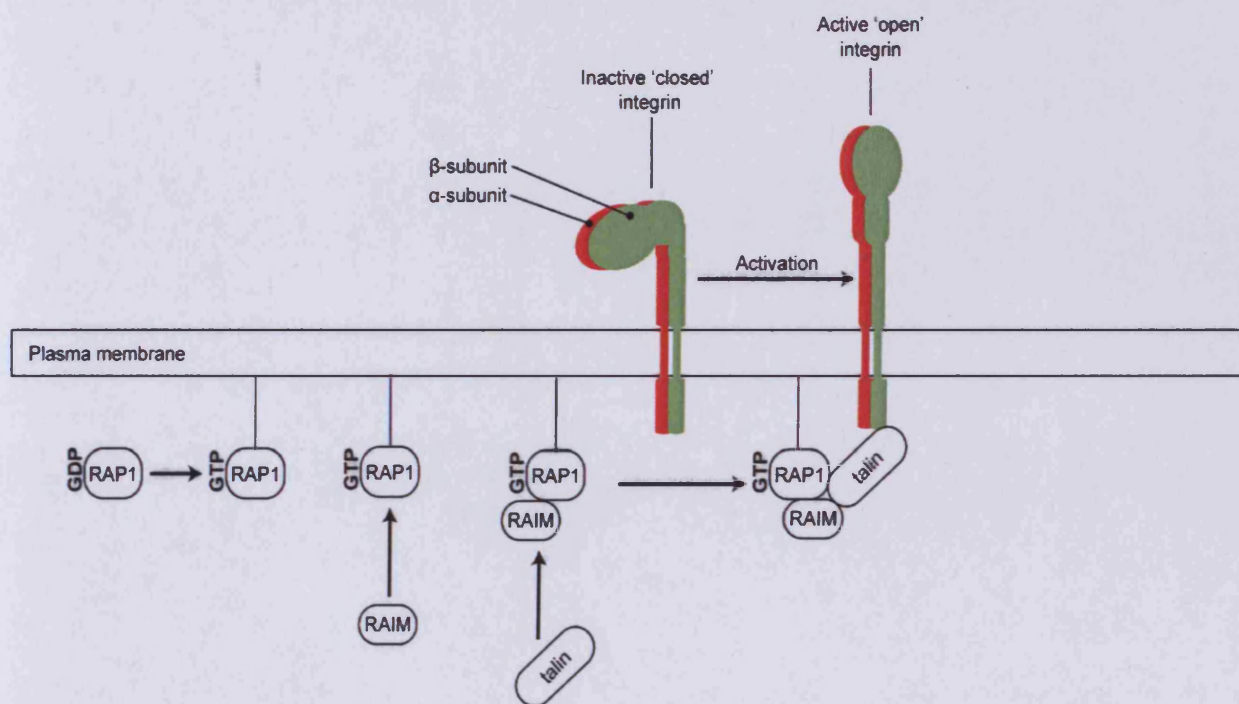


2005). Integrins can undergo conformational changes and clustering which can influence ligand binding. This is important for functionality as conformation of the heterodimers and clustering of heterodimers into oligomers is likely to be dependent on the integrin, cell type and biological context (Shattil et al. 2010). In the majority of integrins, the amino-terminal domain of  $\alpha$ - and  $\beta$ - subunits form a non-covalent bond to form a globular head and a ligand binding site. The ectodomains can exist in a bent 'closed' conformation, an intermediate extended conformation with a 'closed' head and an extended 'open' conformation.

Integrin clustering is defined as the interaction of heterodimers to form hetero-oligomers. This action is integral for outside-in signalling, integrin recycling formation of cell-matrix adhesive structures. Clustering may be caused by the inside-out signalling which recruits protein complexes to the integrin cytoplasmic domains. It can also be caused by the binding of multivalent extracellular ligands to integrin ectodomains or the release of integrins leading to free diffusion of integrins.

The globular head of the extracellular domains of integrin bind ECM components, and interestingly the subunits have no catalytic kinase activity. Instead, signals are transduced through the binding of effector proteins, some of which link directly with the actin cytoskeleton. Together, the integrins with the effector proteins form a multi-protein complexes forming dynamic cell-matrix adhesions (Wiesner, Legate et al. 2005). These adhesions are formulated around activated integrin clusters, and these are essential for cell motility and migration. Adhesion structures must form, disassemble or mature at the leading edge of the cell and disassemble at the rear of the cell for the cell to conduct cell migration.

Inside-out activation of integrins is mediated by the binding of talin to the cytoplasmic domain of the  $\beta$ -integrin subunit, and is a crucial step in integrin  $\beta 1$  and integrin  $\beta 3$  activation (Tadokoro, Shattil et al. 2003). Talin also links integrins to Filamentous actin (F-actin) and other actin binding proteins. This provides the link between the actin cytoskeleton of the cell and the ECM. The talin-integrin interaction is mediated by RAP1, a Ras subfamily member, which is activated by the RAP1 effector protein, RAIM (Boettner and Van Aelst 2009) (Figure 21).



**Figure 21.** Schematic diagram of integrin activation. RAP1 is converted from its inactive GDP state to an active GTP state. Activation of RAP1 leads to the recruitment of RAIM and its binding partner talin to the plasma membrane.

### *B) Cell-matrix interactions*

Cell-matrix adhesive structures can be characterized based on their subcellular location, size and composition. There are 4 types of clearly defined adhesive structures which are focal complexes, focal adhesions, fibrillar adhesions and podosomes. Focal complexes contain talin and vinculin and are small, dot like structures concentrated at membrane protrusions. These either turn-over or mature into focal adhesions. Maturation into focal adhesions occurs with the recruitment of FAK, paxillin,  $\alpha$ -actinin, Src, p130Cas and Crk to the complex (Galbraith, Yamada et al. 2002). These dynamic structures in a 2-Dimensional surface are found at the cell periphery and in more central areas and act as attachment points for actin stress fibres (Wozniak, Modzelewska et al. 2004). They display a slower turn-over rate compared to focal complexes and are associated with less motile cells. Fibrillar adhesions are generally more elongated than focal adhesions and specifically contain  $\alpha 5\beta 1$  integrin and tensin (Pankov, Cukierman et al. 2000). These adhesive structures arrange extracellular fibronectin in a parallel formation with actin stress fibres.

For cell migration to occur the cell must constantly remodel focal complexes into focal adhesions and back again to effectively migrate. This is a dynamic process which involves more than 50 cytoplasmic proteins which present in cell-matrix adhesion structures. Integrins form the link between the ECM and the actin cytoskeleton. They do not possess enzymatic activity and so they recruit signalling proteins to control adhesion-dependent processes. These proteins include integrin binding proteins, adaptor and/or scaffold proteins and enzymes.

Talin is considered to be integral in coupling integrins to the actin cytoskeleton. It is also important in focal complexes/adhesions. Talin self-associates to form a dimer, where the tail region of talin contains binding domains for  $\beta 1$ ,  $\beta 2$ ,  $\beta 3$ ,  $\beta 5$  and  $\beta 7$  integrins (Pfaff, Liu et al. 1998). Talin also contains binding domains for Focal Adhesion Kinase (FAK), non integrin transmembrane receptors, vinculin and actin (Wiesner, Legate et al. 2005). Vinculin is recruited to the early focal adhesion complexes causing an enlargement of focal adhesions by inducing and stabilising 'open' conformation integrin clustering (Wiesner, Legate et al. 2005). The talin-vinculin interaction is integral for vinculin activation (Ziegler, Liddington et al. 2006). Cells with depleted amounts of vinculin have decreased adhesion and increased migratory ability (Saunders, Holt et al. 2006). Vinculin functions as an adaptor protein at focal adhesions. It comprises of 3 domains, an N-terminal head, a flexible proline rich hinge region and a C-terminal tail domain (Humphries, Wang et al. 2007). Paxillin is recruited to the focal adhesion by binding to the vinculin C-terminal tail domain. Like vinculin, it functions as an adaptor protein recruiting other signalling proteins to the focal adhesion. Paxillin may also be recruited by other proteins including parvin, Focal Adhesion Kinase (FAK) and Integrin-Linked Kinase.

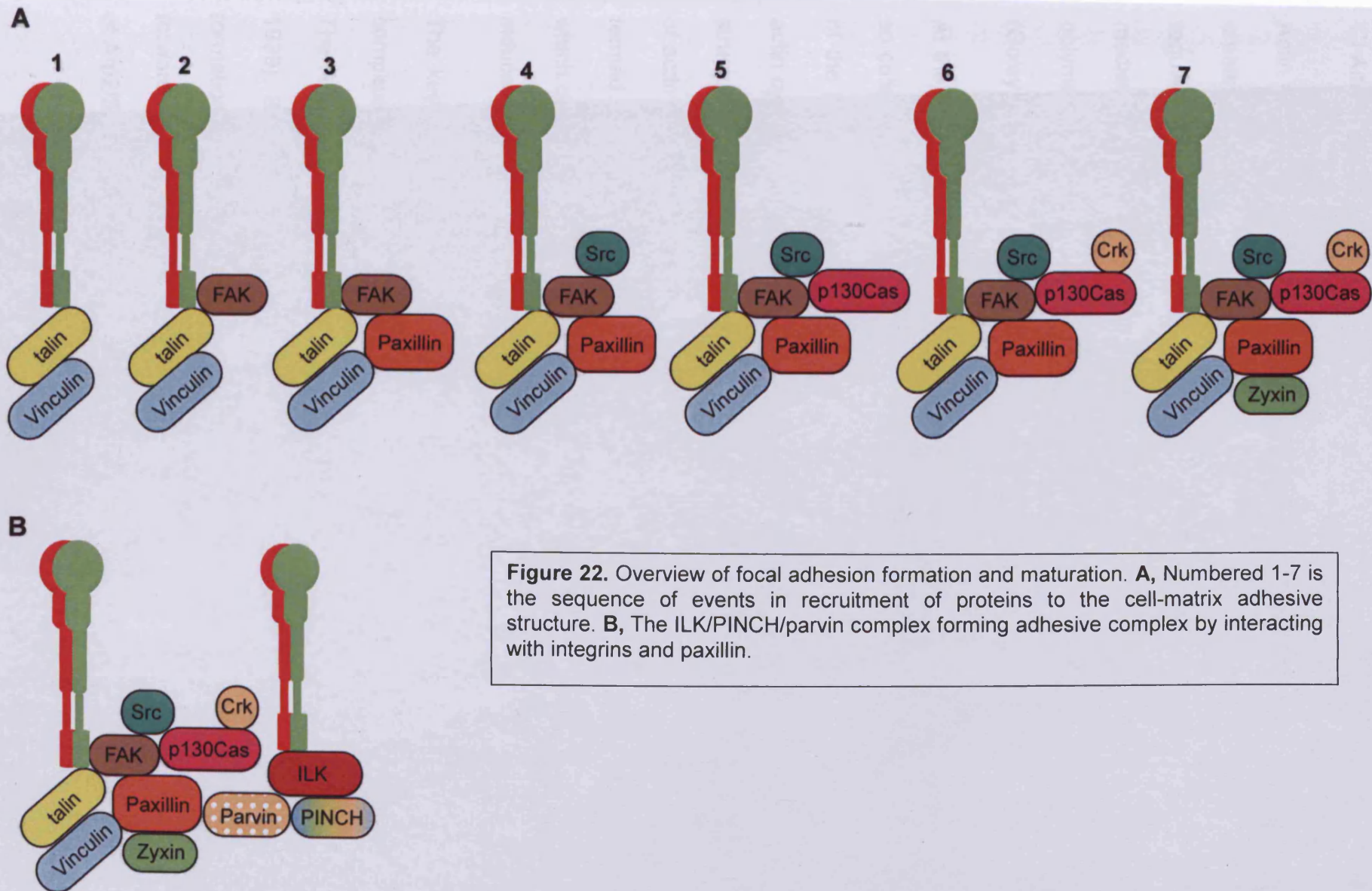
Present in the focal adhesion complex, FAK functions as a tyrosine kinase (Brakebusch and Fassler 2003). Since FAK can strongly bind paxillin and talin, it indicates that the interaction with integrins may occur in an indirect manner (Hildebrand, Schaller et al. 1995). However, co-precipitation experiments have revealed that FAK can bind to the integrin  $\beta 1$  tail region in a direct manner (Schaller, Otey et al. 1995).

The autophosphorylation of FAK at Y397 is induced by integrin activation and this creates a binding site for the Src family kinases. This action causes the recruitment

of another adaptor protein, p130Cas. Its interaction occurs via a SH3 domain which is mediated by phosphorylation by FAK and Src. Tyrosine phosphorylation of p130Cas is an integral event for the recruitment of additional molecules and the release of docking sites for Crk and Nck. The Crk molecule is phosphorylated at T221 by p130Cas, an important event for its localisation to the focal adhesion and for Rac activation. The interaction of p130Cas and FAK with Src is crucial for the targeting of Src to the focal adhesion. Src mediated phosphorylation of Y861 and Y925 creates docking sites for Grb2 which is an important molecule in the MAPK signalling pathway. Src can phosphorylate FAK at Y925 which induces focal adhesion turnover (Frame, Fincham et al. 2002). Focal adhesions mature further with the recruitment of zyxin, an actin scaffold protein (Figure 22A). The zyxin protein is only present in late focal adhesions and it considered a differentiating marker of focal complexes from focal adhesions (Hirata, Tatsumi et al. 2008).

Integrin-linked kinase (ILK) is a non-receptor serine/threonine kinase with a principle role as an adaptor protein. Its principle binding partner is PINCH (Chiswell, Zhang et al. 2008). The ILK-PINCH complex binds to parvin, to form the ILK/PINCH/parvin complex. There are 3 isoforms of parvin, however the  $\alpha$ -parvin isoform is the only parvin to be expressed ubiquitously (Wickstrom, Lange et al.). ILK is able to bind  $\beta$ 1 and  $\beta$ 3 tails of integrins. Downstream of integrins, the parvins regulate cell spreading and actin cytoskeleton organisation. The  $\alpha$ -parvin can bind paxillin at focal adhesions (Yoshimi, Yamaji et al. 2006; Lorenz, Vakonakis et al. 2008). The ILK/PINCH/parvin interaction with integrins enhances the physical interaction of the integrin with the ECM. This complex has an important signalling role and can activate the Akt pathway in a number of cell types (Figure 22B).

Mutational analysis of ILK showed a disruption of the  $\alpha$ -parvin-ILK interaction, which subsequently disrupted the localisation of FAK to focal adhesions (Attwell, Mills et al. 2003). The functionality of ILK has also been highlighted by ILK-null fibroblasts which exhibit reduced cell spreading and migration due to delayed formation of focal adhesions, and a poorly organised actin cytoskeleton (Sakai, Li et al. 2003).



### *C) Actin filaments forming the basis of cell movement*

Actin filaments function as structural scaffolds and as motors for cell movement by creating forward thrust. The assembly and disassembly of actin filaments are regulated by over a hundred actin binding proteins. Actin binding proteins can mediate actin filament dynamics by nucleating, capping, stabilizing, severing, polymerisation, depolymerisation, cross-linking, bundling and sequestering actin (Goley et al. 2006).

At the cell front, the actin assembly drives the extension of membrane protrusions, so called lamellipodia and finger like projections called filopodia. At the leading edge of the lamellipodium adhesive contacts are formed which connect the ECM to the actin cytoskeleton. The anchoring of actin protrusions to the ECM through adhesive structures results in the cell body traction. Forward motion requires the polarisation of actin filaments at the leading edge and depolarisation at the rear. This process is termed actin treadmilling. Actin filaments can grow by the process of polymerisation, which occurs when the actin end is exposed to monomeric actin, and depolarisation induces actin filament stabilisation of actin structures.

The key molecule controlling the formation of the actin protrusions is the Arp2/3 complex which induces branching of actin filaments (Mullins, Heuser et al. 1998). The Arp2/3 complex forms the machinery required for actin-based motility (Welch 1999). Branching of actin filaments at the cell front induces new actin filament formation producing an actin filament network which pulls the plasma membrane forward. The Arp2/3 complex localises only to the leading edge of cells and inhibition of Arp2/3 leads to attenuation of lamellipodia formation (Goley and Welch 2006).



Cofilin (also called ADF) is integral to this process and is localised at the lamellipodium, but not at the leading edge (Svitkina and Borisy 1999). Upon polymerisation, actin hydrolyses ATP creating a difference in concentration at the ATP-bound barbed end of the filament and the ADP-bound pointed end. Cofilin binds the ADP bound actin filaments altering the structure resulting in an increase in the pointed end depolymerisation (Carlier, Laurent et al. 1997). This depolymerisation causes an increase in the concentration of monomeric actin and stabilisation of actin structures.

The small GTPases Rho, Rac and cdc42 are prime regulators of actin filament based movement. They act as molecular switches which cycle between an inactive GDP-bound state and an active GTP-bound state, stimulated by external factors such as growth factors, mechanical stress, integrin signalling and G-protein coupled receptor signalling. Cycling between inactive GDP-bound to active GTP-bound states is induced by GEFs and GTPase-activating proteins (GAPs).

The action of Rho, Rac and cdc42 are tightly regulated in conjunction so that actin polarisation occurs in a controlled manner. Both cdc42 and Rac mediate movement at the leading edge of the cell, by regulating the formation of membrane protrusions whilst Rho regulates the rear of the cell by stimulating the formation of stress fibers critical for actin-myosin mediated cell contraction (Wu, Suetsugu et al. 2004). The formation of stress fibres associated with focal adhesions induces retraction of the rear of the cell. There is interplay between stress fibres and adhesive structures where the formation of focal adhesions initiate the elongation of stress fibres. The tension generated by the stress fibres induces growth of focal adhesions. RhoA activates ROCK which in turn phosphorylates myosin II light chain (MLC). Phosphorylation of MLC is elevated by inhibition of MLC phosphatase by ROCK. This

action assembles the myosin II into myosin filaments and interacts with actin filaments to form the actin contractile stress fibres (BurrIDGE and Wennerberg 2004).

The Arp2/3 complex regulates actin protrusions at the cell front. This complex is activated by the WASP family of proteins. The WASP family is categorised into three groups, WASPs (WASP and N-WASP), WAVEs (WAVE1-3), WASH, WHAMM and JMP (Insall and Machesky 2009). The WASP protein was first to be identified in Wiscott-Alrich Syndrome, and is restricted to the haemopoietic system (Ochs and Thrasher 2006). N-WASP is the near ubiquitously expressed homolog of WASP which functions downstream of cdc42 and Rac (Rohatgi, Ma et al. 1999). N-WASP is localised to the leading edge of lamellipodia in carcinoma cells (Lorenz, Yamaguchi et al. 2004), highlighting its role in cell migration. Also functioning downstream of Rac1, the WAVE proteins are also important activators of the Arp2/3 complex since they are localised at the leading edge of the cell (Insall and Machesky 2009). Whereas N-WASP is regulated by an autoinhibition mechanism, the WAVE proteins are fully active, and their main function is to activate the Arp2/3 complex during membrane protrusions and cell migration.

LIM kinase (LIMK) which functions down-stream of Rho and Rac inactivates cofilin by phosphorylation. Indeed, LIMK is essential regulator of Rho dependent stress fibre/focal adhesion formation and Rac dependent lamellipodia formation (Sumi, Matsumoto et al. 1999). Downstream of Rac activation, LIMK is phosphorylated by PAK1 which inactivates cofilin by phosphorylation (Edwards, Sanders et al. 1999). This thereby inactivates the actin depolymerisation process at the lamellipodium.

Downstream of Rho are the Rho kinases (ROCKs) which function as serine threonine kinases. ROCKs exist as two isoforms, ROCKI and ROCKII. The latter is

largely restricted to the brain and muscle tissues (Riento and Ridley 2003). ROCK1 also phosphorylates LIMK to enhance the phosphorylation and subsequent inactivation of cofilin leading to an increase in actin filament production (Maekawa, Ishizaki et al. 1999) (Figure 23).

Cortactin is an actin filament binding protein which is enriched at the leading edge of lamellipodia where it co-localises with actin and Arp2/3 (Urano, Liu et al. 2001). It forms an important link between the actin cytoskeleton and signalling events which regulate cell adhesion and cell migration (Le Clainche and Carlier 2008). The direct binding to Arp2/3 induces conformational changes to facilitate actin nucleation and filament elongation (Welch and Mullins 2002). N-WASP can cooperate with cortactin to stimulate Arp2/3 and actin assembly. The phosphorylation of cortactin enhances the activation of N-WASP. Cortactin does not have a role in the initial formation of lamellipodia but rather has a role in inducing the focal adhesive contacts formulated after lamellipodia formation (Yamaguchi and Condeelis 2007).

Taken together, cell movement on a 2-Dimensional substrate involves actin protrusion at the leading edge, the formation of new focal complexes at the leading edge. This is followed by formation of contractile stress fibres which forms between focal adhesions. This provides the forward thrust required for forward movement. The final step requires the detractorion from the rear, allowing the cell to move in a forward direction. There is a tight interplay between formation of adhesive structures and actin protrusive force (Figure 24).

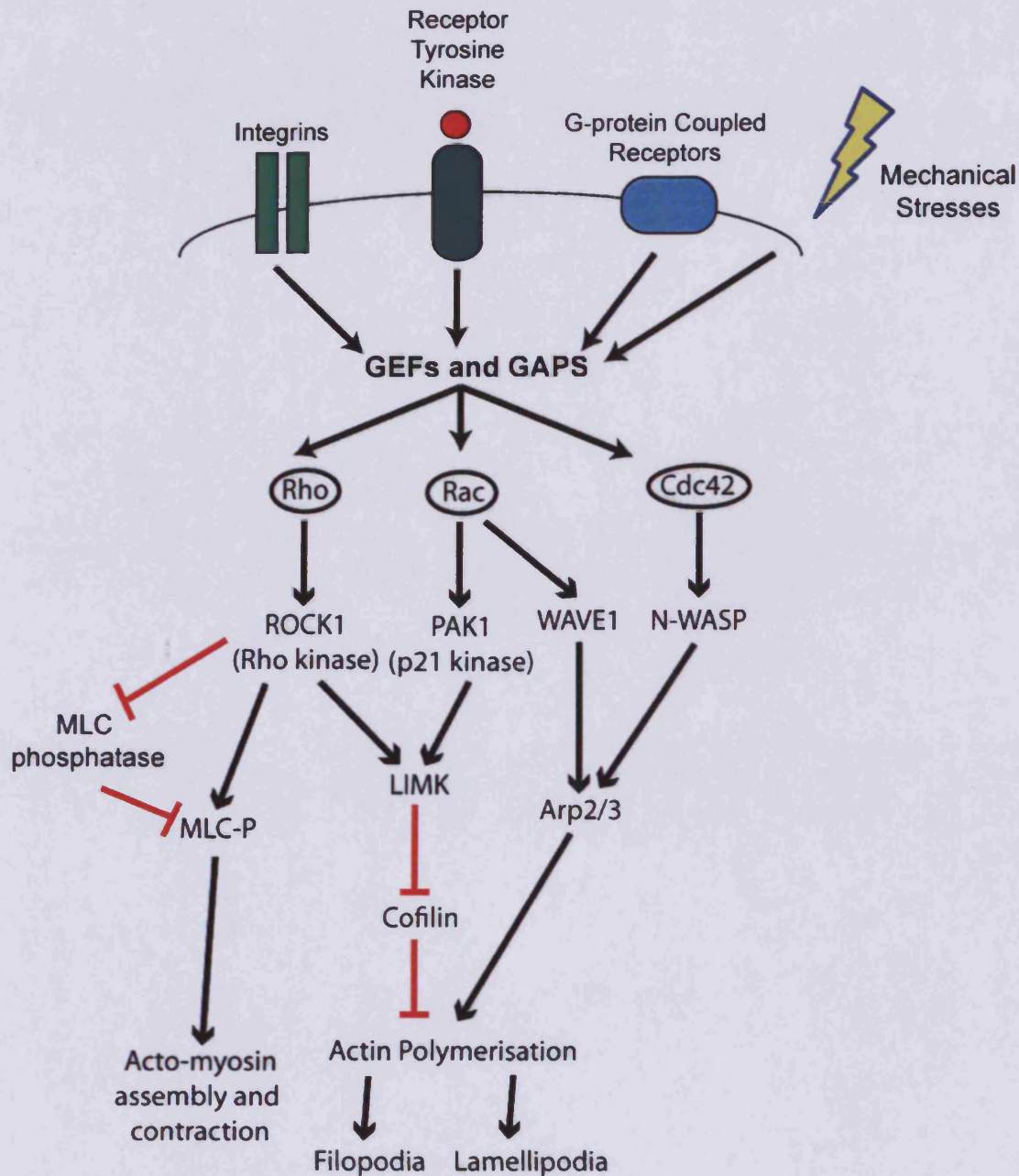
The 2-dimensional study of cell migration and invasion can differ in the morphology, cell-substrate and cell-cell interactions compared to a 3-dimensional environment (Birgersdotter, Sandberg et al. 2005). The importance of 3D ECM has been

highlighted for epithelial cells, where 3D environments promote normal epithelial polarity and differentiation (Roskelley and Bissell 1995).

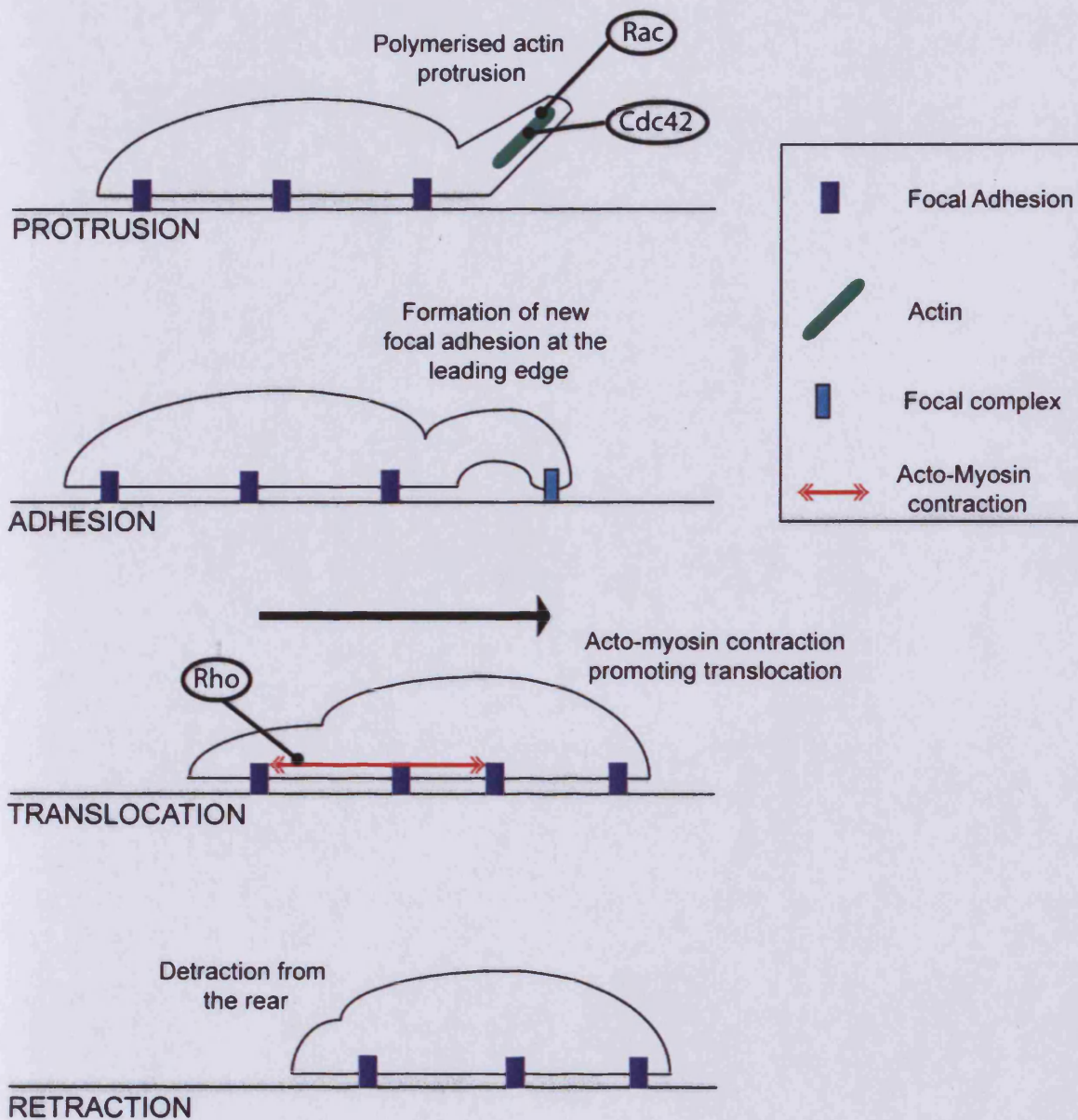
Studies conducted with fibroblasts have shown marked differences between analysing cell migration on 2D substrates and 3D environments. On a 2D substrate fibroblasts display classical focal and fibrillar adhesions. However in 3-dimensional environment, paxillin and  $\alpha_5$  integrin were shown to colocalise to unusual cell-matrix attachments parallel to FN-containing extracellular fibres. These distinctive 3D-matrix adhesions were observed on mesenchymal and migrating mouse neural crest cells. Whereas adhesive structures can form on a 2D-substrate lacking FN, formation of 3D-adhesive structures require FN (Green and Yamada 2007). In addition, the  $\alpha_v\beta_3$  integrin present in 2D fibrillar adhesions is absent from the 3D-matrix adhesions.

Phosphorylation of FAK has been studied in 2D settings as an early step in integrin signaling. In a cell-derived 3D environment, phosphorylation of FAK in cell adhesions is virtually absent from human fibroblasts. Instead, phosphorylation appears to be reduced more than 4-fold and confined to mini focal adhesions (Cukierman, Pankov et al. 2001). The activity level of the Rac GTPase is reduced for fibroblasts within a 3D matrix, giving the cell increased directional movement (Pankov, Endo et al. 2005).

Although 2-dimensional analysis of cell migration is informative, there are clear differences when analysing cell migration in a 3D environment. The 3-dimensional environment allows for greater complexity and may offer greater accuracy to *in vivo* tumour environment. However, effective study of 3D environments requires advanced imaging techniques, and the experimentation of 3-dimensional cell substrates is a relatively novel area.



**Figure 23.** Overview of RhoGTP family signalling during cytoskeletal rearrangement. Rho family of GTPases can be activated by extracellular stimuli signalling through integrins, receptor tyrosine kinases, G-protein coupled receptors or mechanical stress. The stimuli act upon guanine-nucleotide-exchange factors (GEFs) and GTPase-activating proteins (GAPs). Activation of the Rho family of GTPases induces actin polymerisation and formation of actomyosin stress fibres. Red lines indicate inhibition, black lines indicate activation.



**Figure 24.** Overview of cell movement on a 2D substrate. At the leading edge actin protrusions are formed. Protruded actin forms new focal complex adhesions, which stabilise and mature into focal complexes. Acto-myosin contractile forces mediate retraction at the rear of the cell.

#### *D) Cell-cell interactions*

Epithelial architecture is maintained through a number of cell-cell junctions. Adherence and tight junctions form the main adhesive contacts between epithelial cells.

Adherence junctions initiate and stabilise cell-cell contacts, regulate the actin cytoskeleton and intracellular signalling. The adherence junction includes the interactions among the cadherin transmembrane glycoproteins and the catenin family which associate to formulate, maintain and regulate the function of adherence junctions. The 'classical' cadherin family includes E- (epithelial), N- (neural) and P- (placental) cadherin. The structure consists of five extracellular domains, a single transmembrane domain and two cytoplasmic domains. The cytoplasmic domain is highly conserved in the cadherin family, and associates with the catenin family of proteins ( $\alpha$ -,  $\beta$ -, and  $\gamma$ -catenin).

E-cadherin is a prime mediator of epithelial architecture and adhesion. It forms a homodimer by binding another E-cadherin molecule on an adjacent cell. The localisation is specific to specialised zonula adherens-type. The down-regulation of E-cadherin expression in epithelial tumours has been demonstrated. This may occur through a number of mechanisms including mutation, DNA methylation and transcriptional control. The loss of E-cadherin expression has been found in a number of tumours including breast, colon, esophageal, gastric, pancreatic and a number of others cancers. Downregulation of E-cadherin is an essential process during development and involves EMT promoting cell migration (Barbera MJ *et al.* 2004).

$\beta$ -catenin associates with the cytoplasmic domain of E-cadherin, and  $\alpha$ -catenin binds to  $\beta$ -catenin to link the cadherins to the actin cytoskeleton. Cells lacking  $\alpha$ -catenin are unable to form stable adherence junctions despite normal expression of E-cadherin. Cytoplasmic  $\beta$ -catenin is regulated by a protein complex consisting of adenomatous polyposis coli (APC), casein kinase I $\alpha$  (CKI $\alpha$ ), axin and GSK-3 $\beta$  which primes  $\beta$ -catenin for degradation. APC binds to RGS domain of axin and GSK-3 $\beta$  and CKI $\alpha$  interacts with different sites of axin as does  $\beta$ -catenin. Both axin and APC act as scaffold proteins in the destruction complex. This destruction complex acts to degrade  $\beta$ -catenin through phosphorylation at specific amino-terminal Ser and Thr residues (Ser33, Ser37, Thr41 and Ser45) (Mosimann, Hausmann et al. 2009). This leads to ubiquitination of  $\beta$ -catenin promoting its degradation by the ubiquitin-proteasome pathway. The pathway mediates the rapid degradation and elimination of regulatory proteins, and is prevalent within the Wnt signalling pathway. Proteins tagged for proteolysis are ligated by multiple ubiquitin molecules in a multimeric chain (Aberle, Bauer et al. 1997). Disruption of E-cadherin, a cadherin expressed in epithelial cells can alter the cell morphology increasing migratory and invasiveness of cells. Translocation of cytoplasmic  $\beta$ -catenin to the nucleus induces binding to the Tcf/Lcf transcription factor to initiate transcription of numerous genes.

A second catenin, p120-catenin also binds to E-cadherin at the plasma membrane to formulate cell-cell junctions, and is required for increased adhesiveness (Hartsock and Nelson 2008). It also performs a stabilising role preventing internalization of cadherins (Davis, Ireton et al. 2003). It also has a role in cell migration where over-expression leads to formation of actin protrusions and migration in a Rac and cdc42 dependant manner (Noren, Liu et al. 2000) and targeted down-regulation of p12-



catenin leads to attenuated migration and invasiveness (Macpherson, Hooper et al. 2007).

Tight junctions (TJ) are integral to maintain epithelial architecture. They are comprised of several transmembrane proteins and associated intracellular molecules. Perhaps the most extensively studied TJ surface molecule is occludin along with claudins and the junctional adhesion molecule (JAM) families. They also control the passage of ions and solutes between cells (Hartsock and Nelson 2008).

Occludin was one of the first proteins identified as a component of tight junctions. There are two isoforms of occludin; however no differences in structure or function have been reported. Structurally it comprised of 2 extracellular loops, an intracellular turn, N-terminus and C-terminus domains. The C-terminus domain contains serine, threonine and tyrosine residues allowing for phosphorylation events to take place (Sakakibara, Furuse et al. 1997). It can associate with the actin cytoskeleton by binding to actin directly or indirectly with adapter proteins such as ZO and cingulin. The carboxy-terminus cytoplasmic tail of occludin associates with ZO-1, ZO-2 and ZO-3. These ZO proteins interact together to anchor claudins, occludin and JAMS to the actin cytoskeleton. They therefore act as scaffolds for protein interactions at the tight junction. In addition both ZO-1 and ZO-2 have a capacity to bind  $\alpha$ -catenin, and ZO-3 is able to bind p120-catenin. The ZO proteins not only act as linkers to the cytoskeleton, but as scaffold providing a link between tight and adherence junctions.

#### **4.3.2.3 Proteolysis**

##### ***A) Matrix Metalloproteinases***

The ECM is a network of proteins which fill intracellular spaces between cells. Ultimately secreted by cells, the ECM forms a structural scaffold offering mechanical support with participation in a number of cellular functions. Comprised of polysaccharide glycoaminoglycans (GAG) and fibrous proteins the ECM is formed by the secretion of macromolecules. Components of the ECM include collagen, proteoglycans, elastin, fibronectin and laminin. Malignant tumours have a requisite desire to grow, invade and metastasise, and the ECM acts as a barrier against this. To overcome the ECM barrier, transformed malignant cells secrete proteases which degrade the ECM to promote cell invasion and/or tumour angiogenesis. Matrix metalloproteinases represent a major group of enzymes which regulate ECM composition (Visse and Nagase 2003). Classified as matrixins, MMPs are a family of zinc-binding endopeptidases which are known to cleave numerous components of the ECM. MMPs are grouped into a number of categories some of which include, collagenases, gelatinases, stromelysins, matrilysins and membrane bound (Nagase, Visse et al. 2006). 23 MMPs have been identified in humans where all share common structural and functional domains. Structurally MMPs differ depending on which domains are present. All the members have a conserved propeptide domain containing the cysteine switch and a catalytic domain. The catalytic domain contains the zinc binding motif and methionine, with zinc and calcium ions, integral for stability and enzymatic activity (Nagase, Visse et al. 2006). Most MMPs are secreted however; six of the MMPs have transmembrane domains which restrict their expression to the cell membrane. MMPs which are secreted are released latent enzymes and as such require cleavage of the propeptide domain to become active

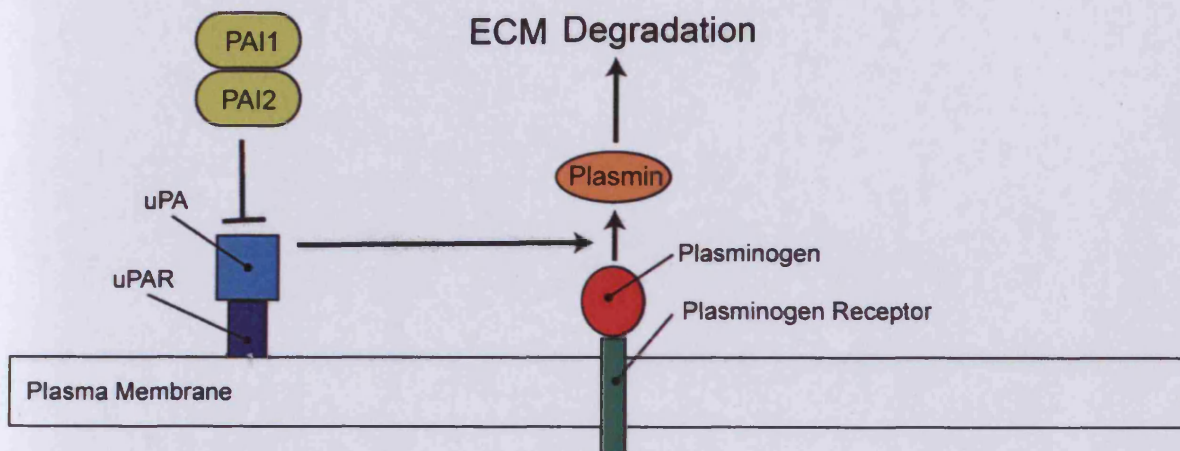
enzymes. Membrane bound MMPs such as MT1-MMP contain a furin-like enzyme recognition motif which allows activation by intracellular proteinases before reaching the cell membrane. For MMP-2 activation, MT1-MMP located at the cell surface is inhibited by the N-terminal domain of Tissue Inhibitor of Metalloproteinases 2 (TIMP-2). The C-terminus of TIMP-2 acts as a receptor for the latent proMMP-2 hemopexin domain. A second adjacent uninhibited MT1-MMP molecule cleaves and activates the tethered proMMP-2 (Itoh and Seiki 2004). The residual propeptide domain left behind following activation is then cleaved by an active MMP-2 molecule. Low expression of TIMP-2 promotes the activation of MMP-2, whilst a higher level of TIMP-2 expression saturates free MT1-MMPs that are required for MMP-2 activation (Strongin, Collier et al. 1995). The gelatinases A and B (MMP-2 and MMP-9 respectively) contain three repeats of a fibronectin type II motif within the catalytic domain which mediate the binding to collagens. Both MMP-2 and MMP-9 cleave gelatin, and type IV, V and XI collagens (Nagase, Visse et al. 2006). The presence and activity of MMPs is increased in malignant cancers with the greatest expression highlighted in invasive areas of the tumour. Several MMPs have been implicated in key molecules associated with tumour invasion, metastasis and angiogenesis. In melanoma, elevated levels of MMP-9 correlated with increased progression, poor survival and metastasis (Deryugina and Quigley 2006). In MMP deficient mice susceptible to carcinogenesis MMP-9 was found to be crucial for the promotion of angiogenesis and tumour vasculature. Within the tumour environment, MMP-2 is upregulated through an OPN dependent manner via interaction with integrin  $\alpha_v\beta_3$  (Rupp, Visconti et al. 2008). In thyroid lesions MMP-2 was readily detected in malignant thyroid lesions (Cho Mar, Eimoto et al. 2006). MMP-9 has additional roles in promoting angiogenesis by regulating the bioavailability of VEGF, a potent mediator

of angiogenesis (Bergers, Brekken et al. 2000). Inhibition of MT1-MMP, MMP-2 and MMP-9 attenuates angiogenesis and lymph node metastasis (Nakamura, Koizumi et al. 2004).

### *B) Urokinase Plasminogen Activator System*

The urokinase plasminogen activator (uPA)/urokinase plasminogen activator receptor (uPAR) system has a role in tumour progression (Andreasen, Egelund et al. 2000). This system comprises of uPA, plasminogen activator inhibitors (PAI's) and uPAR, the G-protein coupled uPA binding receptor. uPA, a serine proteinase is activated by cleavage following the binding to uPAR. The cleavage of uPA is induced by the action of plasmin. uPA generates active plasmin from plasminogen, thus creating a positive feedback loop for uPA activation and plasmin activation (Dass, Ahmad et al. 2008). Plasmin also facilitates the degradation of ECM components including FN, gelatinase, fibrin, laminin and latent collagenases (Fisher, Mackie et al. 2001). This forms the basis of the pro-angiogenic functions of this system. PAI's are anti-proteases which inhibit the action of uPA. PAI-1 binds to active uPA/uPAR complex and induces internalisation of this uPA/uPAR/PAI-1 complex (Figure 25). The internalisation results in the degradation of uPA/PAI-1 by lysosomes, whilst uPAR is recycled and transported back to the cell membrane. In human cancers the plasminogen system has been well established. In epithelial breast cancer, increase in invasiveness was induced in an OPN-uPA dependent manner (Tuck, Hota et al. 2001). OPN upregulates metastatic behaviour through the upregulation of MMP-2 and uPA in an  $\alpha v \beta 3$  integrin dependent manner. The binding of OPN to  $\alpha v \beta 3$  integrin activates Integrin Linked Kinase which induces upregulation of the transcription

factor AP-1 and subsequent upregulation of MMP-2 and uPA. In human mammary epithelial cells where OPN was ectopically expressed resulted in an increase in invasion which was dependant on uPA.



**Figure 25.** The Urokinase Plasminogen Activator System. uPA binds to its receptor, uPAR which activates uPA. Activated uPA generates plasmin from plasminogen to induce ECM degradation.

## **4.4 Aims, Objectives and Experimental Strategy**

Following selection of candidate markers the second major aim of this thesis was to define their role in thyroid tumour invasion. To carry out this aim I focused on 2 genes, Slug and OPN which are over-expressed in PTC. With the use of thyroid cancer cell lines, I investigated the role of these genes in the process of tumour invasion. This effectively defined the processes regulated by these genes, and provided an indication of how they mediate invasion. For this work I utilised thyroid cancer lines, with a view to extending the work to primary models. Using stable shRNA knockdown experiments I conducted experiments to define the role of these markers in tumour invasion. I hypothesised that these selected markers would affect a number of mechanisms.

### **Study the role of Slug and OPN in thyroid tumour invasion**

#### ***1) Define invasive capability in thyroid cancer cell lines and the expression profile of Slug and OPN.***

This work indicates if the expression of Slug or OPN is linked to invasive behaviour.

To do this the following objectives were completed:

1. Characterise invasive capability in a panel of thyroid cancer cell lines
2. Characterise expression of OPN and Slug in a panel of cancer cell lines

## **2) Conduct knock-down of Slug and OPN**

Both Slug and OPN are over-expressed in PTC. To establish the role of these genes in tumour invasion, I hypothesised that shRNA mediated knock-down of expression would confer invasive tumour cells with a non-invasive epithelial phenotype. To achieve this goal the following objectives were carried out:

1. Prepare lentiviruses for the shRNA mediated knock-down of Slug and OPN
2. Conduct the shRNA mediated knockdown of Slug and OPN in a highly invasive cell line over-expressing Slug and OPN
3. Validate the knock-down of Slug and OPN

## **3) Determine if shRNA mediated knock-down of Slug and OPN affects invasive capability**

This would provide evidence that expression of Slug or OPN directly mediates an invasive phenotype. To conduct this aim the following objective was completed:

1. Conduct invasion assays to determine if there is significant reduction in invasive capability.

## **4) Investigate the invasive mechanisms mediated by Slug and OPN which confers an invasive phenotype**

This aim elucidates the mechanisms which Slug and OPN regulate to induce an invasive phenotype. As tumour invasion is a collection of cellular processes, investigations into the roles of extracellular matrix remodelling by proteinases, cell motility, cell-cell adhesion and cell-matrix adhesions were conducted. This provided

direct evidence of Slug and OPN in PTC invasion with an indication of their therapeutic value. To carry out these aims the following objectives were conducted:

*Study of Extracellular matrix remodelling by proteinases*

1. Investigate if the down-regulation of Slug and OPN impacts the urokinase activator system or activity of matrix metalloproteinases.

*Study of cell motility*

1. Investigate how the down-regulation of Slug and OPN impacts the actin based machinery which is required for cell motility
2. Investigate how the down-regulation of Slug and OPN impacts the motile behaviour of cells.
3. Investigate the role of Hepatocyte Growth Factor in motile behaviour, and the regulation by Slug and OPN.
4. Analyse if variant isoforms of CD44 are a contributing factor mediating HGF-dependant motility.

*Study of cell-cell interactions to investigate if the down-regulation of Slug and OPN confer mesenchymal to epithelial transition*

1. Investigate how the down-regulation of Slug and OPN impacts tight junction regulation
2. Investigate if the down-regulation of Slug and OPN regulates adherence junctions



3. Investigate how the down-regulation of Slug and OPN affects cell-matrix interactions and how this relates to motile behaviour.

# **5.0 Materials and Methods**

## **5.1 Cell culture**

For the culture of cell lines, cells were grown in a humidified atmosphere containing 20 % O<sub>2</sub>, 5 % CO<sub>2</sub> at 37°C. Cells were either grown in RMPI, Dulbecco's modified Eagles' medium or the 1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12. All were supplemented with 10% foetal calf serum, Penicillin and Streptomycin with glutamine (Table 5.0).

## **5.2 Preparation of Lentiviral supernatants**

The Lentiviral vectors used in my investigations were the MISSION™ Slug shRNA and MISSION™ OPN shRNA both in a pLKO.1-puro vector (both purchased from Sigma-Aldrich, UK) (Table 5.1). For each transfection sample DNA-Lipofectamine™ 2000 complexes were prepared. This involved diluting 9µg of the ViraPower™ Packaging Mix (Invitrogen, Paisley, UK) and 3µg of shRNA plasmid DNA (12µg total) in 1.5mL of Opti-MEM® medium without serum. This was gently mixed. In a separate sterile 5 mL tube, Lipofectamine™ 2000 was mixed gently before use and 36µL diluted in 1.5mL of Opti-MEM® medium without serum. This was followed by gently mixing and incubating for 5 minutes at room temperature. After a 5 minute incubation, the diluted DNA with the diluted Lipofectamine™ 2000 was combined and gently mixed. This mixture was incubated for 20 minutes at room temperature to allow the DNA-Lipofectamine™ 2000 complexes to form.

Whilst the DNA-Lipofectamine complexes were forming, 293FT cells were trypsinised and resuspended at a density of  $1.2 \times 10^6$  cells/mL in 10mL of growth medium containing serum without antibiotics. The DNA-Lipofectamine™ complexes

were added to a 10cm tissue culture plate containing 5mL of Opti-MEM® Medium with serum and without antibiotics. To this 5mL of the 293FT cell suspension ( $6 \times 10^6$  total cells) was added to the plate. The plate was mixed by rocking the plate back and forth and incubated overnight at 37°C in a humidified 5% CO<sub>2</sub> incubator. The following day the medium was removed and replaced with 10mL complete culture medium without antibiotics. This was left to incubate at 37°C in a humidified 5% CO<sub>2</sub> incubator. Virus-containing supernatants were harvested after 48 hours post-transfection by removing medium into to a 15mL sterile, capped, conical tube. This was centrifuged at 1000g for 15 minutes at 4°C to pellet debris. Viral supernatants were stored at -80°C using cryovials in 1mL aliquots.

Lentiviral infection of cells was conducted according to the materials and methods section in Part I (2.2.3)

Oncogene/plasmid	Producer cells	Selection
SlugshRNApLKO.1-puro	293FT	Puromycin
OPNshRNApLKO.1-puro	293FT	Puromycin

**Table 5.1.** Lentiviral plasmids with producer cells used and corresponding selection

Cells	Source	Growing media	Usage
Primary human thyroid epithelial cells	Normal human thyroid tissue	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Identification of candidate genes using Microarray
ORI-3	Human primary thyroid follicular epithelial cells immortalized by SV40	RPMI medium, supplemented with 10% fetal calf serum	Thyroid negative control for invasion assays Negative control for the study of Slug and OPN in invasion Characterisation of B-Raf isoforms
K1	Human papillary thyroid carcinoma cell line	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms Positive control for tumour invasion Positive control for the study of Slug and OPN in invasion
FTC133	Human follicular thyroid carcinoma cell line	RPMI medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms Study of tumour invasion
FTC236	Human follicular thyroid carcinoma cell line derived from the metastasis of FTC133	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms Study of tumour invasion
Hth74	Human anaplastic thyroid carcinoma cell line	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms Study of tumour invasion
SW1736	Human anaplastic thyroid carcinoma cell line	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms
TT	Human medullary thyroid carcinoma cell line	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium with 15% fetal calf serum.	Characterisation of B-Raf isoforms
K5	Human follicular thyroid carcinoma cell line	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms

UCLA-RO82	Human follicular thyroid carcinoma cell line	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms
BCPAP	Human papillary thyroid carcinoma cell line	1:1 mixture of Dulbecco's modified Eagles' medium and Hams' F12 medium supplemented with 10% fetal calf serum.	Characterisation of B-Raf isoforms Study of tumour invasion
A431	epidermoid carcinoma cell line	RPMI medium, supplemented with 10% fetal calf serum supplemented with 10% fetal calf serum.	Positive control for adherence junctions
MCF7	Human breast carcinoma cell line	RPMI medium, supplemented with 10% fetal calf serum	Control cell line used in the study of Slug and OPN in invasion
MDAMB231	Human breast carcinoma cell line	Dulbecco's modified Eagles' medium with 10% fetal calf serum	Control cell line used in the study of Slug and OPN in invasion
293FT	Human kidney cell line	Dulbecco's modified Eagles' high glucose medium. (supplemented with 10% fetal calf serum depending on experimental conditions)	Production of lentiviruses
Phoenix	Amphotrophic retrovirus packaging cell line	Dulbecco's modified Eagles' medium with 10% fetal calf serum	Production of BRAF <sup>V600E</sup> retroviruses
Ψ-CRIP DOEJ	RASV12 retrovirus producer cell line	Dulbecco's modified Eagles' medium with 10% fetal calf serum	Production of RASV12 retroviruses
Ψ-CRIP RET/PTC1	RET/PTC1 retrovirus producer cell line	Dulbecco's modified Eagles' medium with 10% fetal calf serum	Production of RET/PTC1 retroviruses

**Table 5.0.** Cells with corresponding culture conditions, source and usage.

### **5.3 Cell Lysis**

Cells were grown to 70% confluence and scraped into 1mL STE buffer (10mM Tris pH 8.0, 150mM NaCl, 1mM EDTA pH 8.0) and centrifuged at 1000rpm at 4°C for 5 minutes. After discarding supernatant cells were lysed for 5 minutes on ice in 2.5 times the volume of lysis buffer (150mM NaCl, 50mM Tris, 5mM EDTA and 1% NP40 with 30µL/mL PMSF, 10µL/mL protease inhibitor cocktail (Calbiochem, Nottingham, UK) and 10µL/mL phosphatase inhibitor cocktail (Calbiochem, Nottingham, UK). The lysed cell mixture was centrifuged at 56,000rpm at 4°C for 30 minutes. The supernatant was removed and stored at -20°C for future use.

### **5.4 Protein Quantification**

To determine protein concentration the Bradford method of quantification was utilised. Bovine Serum Albumin (BSA) standards (Pierce Ltd, UK) were prepared in triplicate at concentrations of 100µg/mL, 200µg/mL, 300µg/mL and 400µg/mL by diluting 2mg/mL BSA in 1/10 diluted lysis buffer. Protein samples were diluted 1/10 in water and 40µL of each standard and sample was aliquoted into 1.6mL semi-micro cuvettes to which 1mL of Bradford Coomassie Plus Protein Assay reagent (Thermo Fisher Scientific, Rockford, IL, USA) was added. Using the 1/10 diluted lysis buffer as a reference, the standards and samples were read on a spectrophotometer at a wavelength of 595nm. Protein concentrations were calculated from the colorimetric values obtained by producing a standard curve.

## **5.5 Sodium Dodecyl Sulfate (SDS) Polyacrylamide Gel Electrophoresis (PAGE)**

### **Western Immunoblotting**

Samples were prepared by addition of 3X loading buffer (300 $\mu$ L 0.5M Tris pH 6.8, 300 $\mu$ L Glycerol, 300 $\mu$ L 10% SDS, 60 $\mu$ L 1% bromophenol blue and 30 $\mu$ L 2-mercaptoethanol) followed by boiling for 5 minutes. The Precision Plus Molecular Weight Marker (Bio-Rad, Hertfordshire, UK) was also prepared in the same way. Electrophoresis was performed using an SE-250 mini-vertical electrophoresis unit (GE Healthcare Life Sciences, Buckinghamshire, UK) at 100V until the desired separation had occurred. Prior to the transfer of proteins onto a Polyvinylidene Fluoride (PDVF) membrane (Millipore, Billerica, MA, USA), the membrane was equilibrated in methanol. The transfer of proteins was performed using a TE-22 Mini-tank transfer unit (GE Healthcare Life Sciences, Buckinghamshire, UK) overnight at 10V.

For immunoprobng, the membranes were blocked according to antibody conditions for 4 hours at 4°C on a shaker. Primary antibodies were diluted according to antibody conditions and probed at 4°C overnight on a roller (Table 5.2). Secondary antibodies were either anti-mouse Horseradish Peroxidase, anti-rabbit Horseradish Peroxidase or anti-goat HorseRadish Peroxidase. This was dependant on the antibody being used. Membranes were washed several times with 1XTBS 0.2% Tween20 after antibody incubations. For chemiluminicent detection ECL Plus reagent (GE Healthcare Life Sciences, Buckinghamshire, UK) was used. For detection membranes were ultimately developed onto ECL-film.



## 5.6 Zymogram Analysis

Samples were prepared by growing cells in a T<sub>75</sub> tissue culture flask to 80% confluence, aspirating the medium and replenishing with serum-free medium. These were incubated overnight at 37°C in a humidified 5% CO<sub>2</sub> incubator. The following day, the medium was harvested, centrifuged at 1000rpm for 5 minutes to remove any cell debris and stored at -20°C for future use.

Gelatin 10% Ready Gel Zymogram Gels were commercially obtained (Biorad laboratories). Samples were thawed to which 2x Tris-Glycine SDS buffer was added at a ratio of 1:1. The buffer consisted of 2.5 mL 0.5M Tris-HCl, pH 6.8, 2 mL Glycerol, 4mL 10% (w/v) SDS, 0.5mL 0.1% Bromophenol Blue and H<sub>2</sub>O to 10mL. Samples were applied to the gel using the same conditions as SDS-PAGE. The gel was then incubated in Renaturing Buffer, which consisted of 2.5% Triton-X-100 (v/v) in H<sub>2</sub>O for 30 minutes with gentle agitation. The Renaturing Buffer was decanted and placed in the Developing Buffer (50mM Tris-base, 0.2M NaCl, 5mM CaCl<sub>2</sub>, 0.7mg ZnCl<sub>2</sub>, 0.02% NaN<sub>3</sub> and H<sub>2</sub>O to 1L) for 5 hours at 37°C. The final step consisted of staining with Coomassie Blue R-250 (500mg Coomassie Blue R-250, 225mL methanol, 225mL deionised H<sub>2</sub>O, 50mL Acetic acid) for 30 minutes and destained using methanol : acetic acid : H<sub>2</sub>O (50:10:40) Destaining solution until clear bands were visible.

## **5.7 OPN Enzyme-linked immunosorbent assay (ELISA)**

Human OPN quantikine immunoassay was commercially obtained (R&D Systems). To each well of the micro-plate, 100 $\mu$ L of Assay Diluent RD1-6 was added, followed by the addition of 50 $\mu$ L of standard, control, or sample per well. Each standard, control and sample was analysed in duplicate. This was covered with the supplied adhesive strip. The plate was then incubated for 2 hours at room temperature. Each well of the plate was aspirated and washed by filling each well with Wash Buffer (400 $\mu$ L) using a squirt bottle, with the process repeated three times for a total of four washes. After the last wash, any remaining Wash Buffer was removed by aspiration and decanting, followed by blotting against clean paper towels. To each well, 200 $\mu$ L of OPN conjugate was added and the plate covered by a new adhesive strip. This was incubated for 2 hours at room temperature. Using the Wash Buffer, the wash step was then repeated. To each well 200 $\mu$ L of substrate solution was added, which was then incubated for 30 minutes at room temperature in the dark. To stop the reaction, 50 $\mu$ L of Stop Solution was added to each well. The optical density of each well was determined within 30 minutes using a FLUOstar OPTIMA microplate reader (BMG Labtech, Germany) set to 450nm. Readings at 540 nm were also determined and these were subtracted from the readings at 450 nm. This subtraction corrected the optical imperfections in the plate.

## **5.8 Immunofluorescence**

Cells were grown on glass coverslips (RA Lamb, Lauda-Konlgshofen, Germany) and fixed in 4% paraformaldehyde and rinsed three times with 1XPBS. Paraformaldehyde fixed cells were permabilised with 0.2% Triton-X-100 for 20 minutes at room temperature, and were blocked in goat serum (1:20 in 1XPBS) for 1 hour. Primary antibodies were incubated overnight at 4°C. All antibodies were diluted and washed with 1XPBS 0.001% Tween20. Secondary antibodies were diluted in PBS and incubated at room temperature for 1 hour. Coverslips were counterstained with 20µL DRAQ5 (Biostatus Ltd, Shepshed, UK) or DAPI nuclear stain for 10 minutes and mounted on glass slides using Prolong Gold Antifade reagent (Invitrogen, Paisley, UK) and sealed with clear nail varnish.

Details regarding antibodies, company, and concentrations are detailed in Table 5.2.

## **5.9 Phalloidin Staining**

Cells grown on coverslips were washed twice with 1xPBS. Cells were fixed with 4% paraformaldehyde and washed 3 times with 1xPBS. Cells were permeabilised with 0.1% Triton-X-100 for 20 minutes and rinsed 3 times with 1xPBS. Cells were incubated for 20 minutes in the dark with a phalloidin-FITC (Sigma-Aldrich, UK) antibody (50µg/mL). Coverslips were washed with 1xPBS and counterstained with DAPI for 10 minutes. Coverslips were mounted on glass slides using Prolong Gold Antifade reagent (Invitrogen, Paisley, UK) and sealed with clear nail varnish.

Application	Protein name	1 <sup>o</sup> Antibody	Antibody source	Specificity	Conc.	Blocking solution	2 <sup>o</sup> Antibody	Conc.
<b>IF</b>	Phospho-Cortactin	(Y466)	Abcam	H, M	1:500	10%FCS/PBS	Alexa Fluor 568	1:500
	E-cadherin	(HECD-1)	Abcam	H	1:100	10%FCS/PBS	Alexa Fluor 488	1:500
	Vinculin	(sc73614)	Santa-Cruz Biotechnology Inc	H	1:100	10%FCS/PBS	Alexa Fluor 488	1:500
	WAVE1		Gift from Prof. W. Jiang	H	1:250	10%FCS/PBS	Alexa Fluor 568	1:500
	N-WASP		Gift from Prof. W. Jiang	H	1:250	10%FCS/PBS	Alexa Fluor 568	1:500
	Slug	(G18)	Santa-Cruz Biotechnology Inc	H	1:200	10%FCS/PBS	Alexa Fluor 488	1:500
	$\beta$ -catenin	(610154)	BD Transduction Laboratories	H, M	1:50	10%FCS/PBS	Alexa Fluor 568	1:500
	Pan-cadherin	(CH-19)	Abcam	H, M	1:100	5%FCS/PBS	Alexa Fluor 488	1:500
	Integrin $\beta$ 3	(BV4)	Abcam	H	1:100	5%FCS/PBS	Alexa Fluor 488	1:500
	CD44	(156-3C11)	Cell signalling Technology	H	1:400	20%FCS/PBS/ Triton-X	Alexa Fluor 568	1:500
<b>WB</b>	Snail	(L70G2)	Cell signalling Technology	H, M	1:1000	5% milk/1xTBS/0.1% Tween	RAM	1:2000
	Slug	(G18)	Santa-Cruz Biotechnology Inc	H	1:1000	5% milk/1xTBS/0.1% Tween	RAG	1:2000
	B-Raf	(24462)	Upstate Biotechnology	H	1 $\mu$ g/ml	5% milk/1xTBS/0.1% Tween	GAR	1:2000
	$\beta$ -catenin	(610154)	BD Transduction Laboratories	H, M	0.25 $\mu$ g/ml	5% BSA/1xTBS/0.1% Tween	RAM	1:2000
	Occludin	(33-150)	Invitrogen	H, M	1 $\mu$ g/ml	5% milk/1xTBS/0.1% Tween	GAR	1:2000
	MT1-MMP	(EP1264Y)	Abcam	H, M	1:2000	5% BSA/1xTBS/0.1% Tween	RAM	1:2000
	Actin	(048K4861)	Sigma-Aldrich	H	1:750	5% milk/1xTBS/0.1% Tween	RAG	1:2000

**Table 5.2.** Antibodies used for Immunofluorescence (IF) and Western Blot analysis (WB). (RAM; Rabbit anti-mouse, GAR; Goat anti-rabbit, RAG; Rabbit anti-goat, H; Human, M; Mouse). WAVE1 and N-WASP antibodies were a gift from Prof. Wen Jiang (Department of Surgery, Anaesthetics, Obstetrics & Gynaecology, Cardiff University)

### **5.10 Mini-preparation of plasmid DNA**

MISSION™ Bacterial glycerol stocks containing Slug shRNA or OPN shRNA in a pLKO.1-puro vector were commercially obtained (Sigma-Aldrich, UK). The glycerol stocks were streaked onto agar-carbenicillin (200µL carb/100mL LB) plates and left to incubate at 37°C overnight. Colonies were picked and inoculated in 5mL LB-carbenicillin and incubated at 37°C overnight with agitation. Mini-preps were conducted of the inoculated LB were conducted using the PureLink™ Quick Plasmid Miniprep Kit (Invitrogen, Paisley, UK).

The bacterial culture was pelleted and the medium removed from the cell pellet. This was completely resuspended in 250µL Resuspension Buffer (R3) with RNase A making sure no cell clumps remained. To the resuspended bacterial pellet, 250µL Lysis Buffer (L7) was added to cells and mix gently by inverting the capped tube 5 times. This was followed by a 5 minute incubation at room temperature. After the incubation, 350µL Precipitation Buffer (N4) was added. This was mixed immediately by inverting the tube until the solution was homogeneous. The mixture was centrifuged at 12,000g for 10 minutes at room temperature using a microcentrifuge to clarify the lysate from lysis debris. The supernatant was then loaded onto a spin column. The Spin Column with supernatant was placed into a 2mL Wash Tube. This was centrifuged at 12,000g for 1 minute. The flow-through was discarded and the column placed back into the Wash Tube. Next, 500µL Wash Buffer (W10) with ethanol was added to the column and incubated for 1 minute at room temperature. This was then centrifuged at 12,000g for 1 minute. The flow-through was again discarded and the column placed back into the

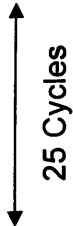
Wash Tube. Following this step, 700 $\mu$ L Wash Buffer (W9) with ethanol was added to the column. The column was centrifuged at 12,000g for 1 minute. The flow-through was discarded and the column placed back into the Wash Tube. To remove any residual Wash Buffer (W9), the column was centrifuged at 12,000g for 1 minute and the flow-through discarded. The Spin Column was placed in a clean 1.5mL Recovery Tube to which 50 $\mu$ L of TE Buffer (TE) was added to the centre of the column. The column was incubated for 1 minute at room temperature, followed by centrifugation at 12,000g for 2 minutes. The Recovery Tube contained the purified plasmid DNA.

### **5.11 cDNA synthesis and Reverse Transcription Polymerase Chain Reaction (RT-PCR)**

Following RNA extraction, 1µg of RNA was combined with 2µL (1µg/µL) Random Primers (Invitrogen, Paisley, UK) and 8µL sterile water and incubated for 10 minutes at 70°C. To this 4µL 5X buffer, 2µL 0.1M DTT, 1µL 10m dNTP, 1µL Superscript II and 1µL RNaseOUT™ (All from Invitrogen, Paisley, UK) were added and incubated for 1 hour at 42°C. This was followed by conventional PCR.

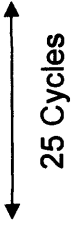
Primers dedicated to each target were designed and ordered through a commercial company (MWG-Biotech). PCR program was conducted according to Table 5.3. GAPDH primers which were commercially obtained (Biomol, Germany). The PCR program was conducted according to Table 5.4.

For the conventional PCR, 2µL of the cDNA, 8.5µL H<sub>2</sub>O, 5µL 10x Buffer, 1µL dNTPs, 0.5µL Taq polymerase, 3µL Magnesium Chloride (all from Thermo-Fisher, UK) with 2.5µL of forward and reverse primer (10 pmol/µL) was used. Primer sequences are detailed in Table 5.5.

PCR Step	Program	Temperature	Time	Number of Cycles
Initial denaturing		94°C	2 minutes	 25 Cycles
Denaturing		94°C	15 seconds	
Annealing		Primer dependant	30 seconds	
Elongation		72°C	3 minutes	
Final Elongation		72°C	10 minutes	

**MWG primers**

**Table 5.3.** PCR program for primers ordered through MWG. Annealing temperature is primer dependant.

PCR Step	Program	Temperature	Time	Number of Cycles
Initial denaturing		94°C	2 minutes	 25 Cycles
Denaturing		94°C	15 seconds	
Annealing		55°C	30 seconds	
Elongation		72°C	3 minutes	
Final Elongation		72°C	10 minutes	

**GAPDH primers**

**Table 5.4.** PCR program for GAPDH primers ordered from Biomol. PCR program was conducted according to manufacturer instructions.



	Length (bp)	Primer Sequences (5'-3'; a, forward; b, reverse)	Cycles	Annealing Tempertaure (°C)
B-Raf exon 8 -18	1598	a- CCAATTGGTTGGGACACTG b- GTTCTGATGCACTGCGGTG	25	56
B-Raf exon 10 – 18	922	a- GCCTCATTACCTGGCTCAC b- GTTCTGATGCACTGCGGTG	25	56
E-cadherin	612	a- TGGGCTGGACCGAGAGAGTT b- ATCTCCAGCCAGTTGGCAGT	25	55
Osteopontin	a-450 b-408 c-369	a- AGCCAAACGCCGACCAAG b- GGTCCGTGGGAAAATCAGTG	25	56
Slug	258	a- AGATGCATATTCGGACCCAC b- CCTCATGTTTGTGCAGGAGA	25	57
uPA	199	a- ACTACTACGGCTCTGAAGTCACCA b- GAAGTGTGAGACTCTCGTGTAGAC	25	50
uPAR	157	a- CTGGAGCTTGAAAATCTGCCG b- GGTTCCTCGGTTTCGTGAGTGC	25	50
MT1-MMP	221	a- GGATACCCAATGCCATTGGCCA b- CCATTGGGCATCCAGAAGAGAGC	25	51
MMP-2	225	a- AGATCTTCTTCTTCAAGGACCGGTT b- GGCTGGTCAGTGGCTTGGGGTA	25	51
CD44s	157	a- AAAGGAGCAGCACTTCAGGA b- TGTGTCTTGGTCTCTGGTAGC	25	51
CD44v6	175	a- AACTGATATTCTTCTCACAG b- CTTGTTAAACCATCCATTACCAG	25	50
CD44v7	200	a- CTCAAACTGCATGGTCACAG b- AAAATCTCAGAGGCTATTAC	25	50
CD44v10	310	a- CTGATTCCACCTCCACACAG b- GATAATAAATGCCAAATTACCTG	25	50
GAPDH	240	a- TGATGACATCAAGAAGGTGGTGAAG b- TCCTTGGAGGCCATGTGGGCCAT	25	55

**Table 5.5.** Primer sequences detailing product size, number of cycles and melting temperature.

# 6.0 Results

## **6.1 The role of Slug and OPN in Thyroid Tumour Invasion**

The differential diagnosis between follicular adenoma and follicular carcinoma is difficult to determine following major histological examination after thyroidectomy (Davis, Gordon et al. 1991). The identification of additional biomarkers of invasion would greatly facilitate in the differential diagnosis, and further provide opportunities for identifying new therapeutic targets.

Our laboratory has established a primary model representing FA, and two additional models representing PTC. Using these I conducted a Microarray gene profiling experiment to identify novel candidate genes which may regulate an invasive phenotype. In parallel, I had conducted a literature analysis to identify additional biomarkers. From this analysis I identified two genes, Slug and Osteopontin (OPN) which are implicated in thyroid cancer.

Osteopontin, a secreted phosphoprotein regulates numerous processes pertaining to tumour invasion and metastasis. OPN is overexpressed and enhances PTC invasiveness in RET/PTC transformed rat thyroid cells and correlates with aggressive clinicopathological features of PTC (Guarino, Faviana et al. 2005). OPN-integrin binding directly mediates migration and invasion of tumour cells, and is known to enhance thyroid papillary carcinoma invasiveness. This includes increasing cell migration, motility and extracellular matrix degradation. The mechanistic explanation of how OPN regulates thyroid tumour invasion requires further investigation. The overexpression of OPN and its implication in invasion has already been elucidated. However, the precise mechanisms regulating the invasive program mediated by OPN in thyroid cancer requires further investigation.

The Slug transcription factor has been implicated in papillary carcinoma. Down-regulation of E-cadherin mediated by Slug has been demonstrated in thyroid carcinoma (Hardy, Vicente-Duenas et al. 2007). However, it is unclear if other invasive processes are affected by the overexpression of Slug.

We therefore hypothesised that OPN and Slug may regulate multiple invasive processes. I therefore aimed to assess upstream regulators and downstream effectors of these nominated candidates. To achieve this overall aim, lentiviral mediated shRNA knockdown of Slug and OPN was used.

## 6.2 Characterising the invasive behaviour across a panel of thyroid cancer cell lines.

To investigate invasive behaviour *in vitro*, I used a modified *in vitro* boyden chamber invasion assay utilising reconstituted growth factor reduced basement membrane matrix. The reconstituted basement membrane matrix acts as a barrier which invasive cells have to pass through. This effectively reflects the *in vivo* situation of tumour invasion. The cell lines analysed in this study were from different thyroid tumours (Table 6.0) and the invasive potential of each cell line was analysed. Determining the thyroid cell lines which display the greatest degree of invasive behaviour was important aim as it establishes the experimental controls to assess possible roles of Slug and OPN in tumour invasion.

Cell line	Source
ORI-3	Normal epithelium transformed with SV40
K1	Human papillary thyroid carcinoma cell line
BCPAP	Human papillary thyroid carcinoma cell line
FTC133	Human follicular thyroid carcinoma cell line derived from the primary tumour of a patient
FTC236	Human follicular thyroid carcinoma cell line derived from the metastasis of the primary tumour (FTC133)
Hth74	Human anaplastic thyroid carcinoma cell line

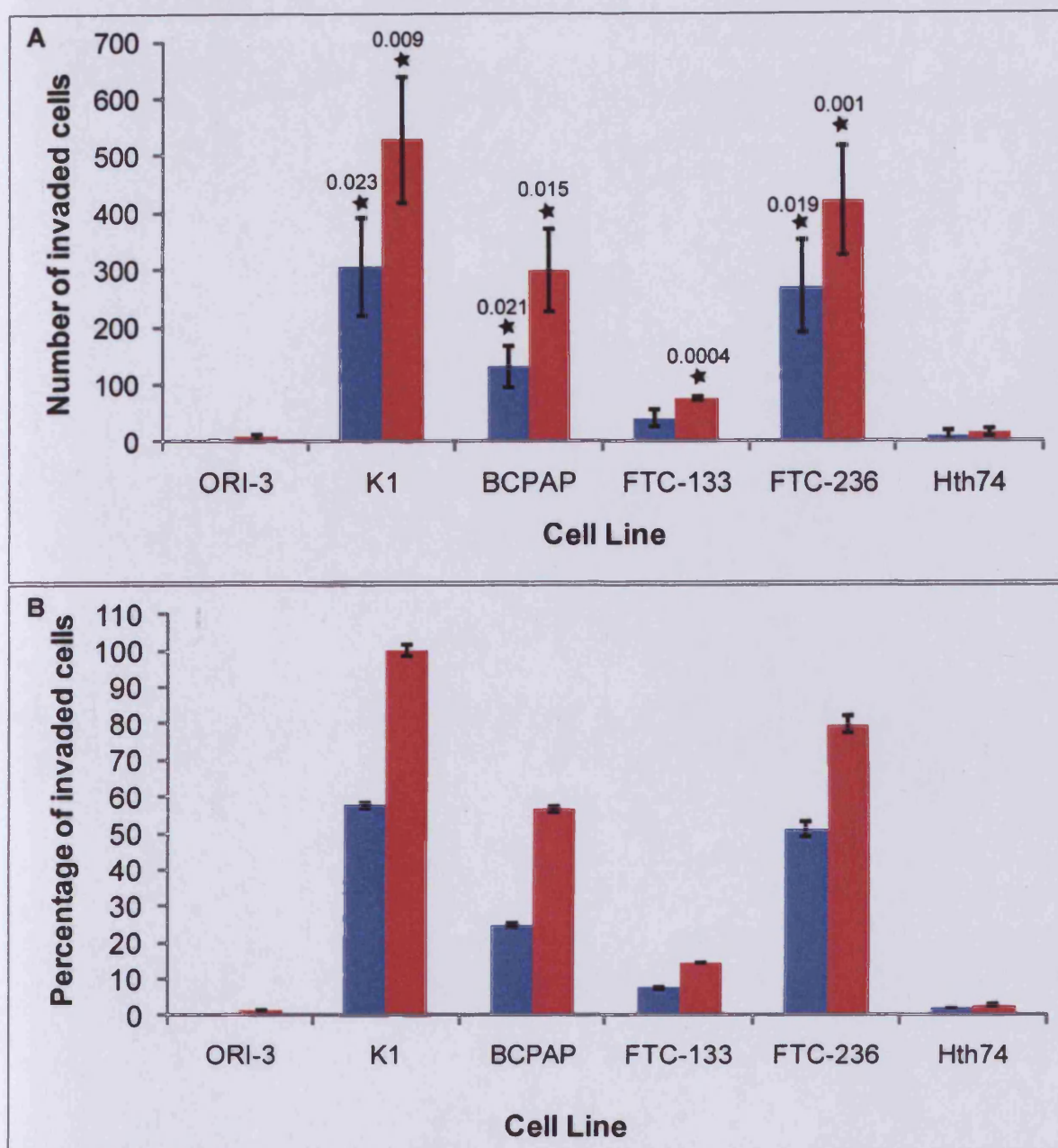
**Table 6.0.** The thyroid cell lines and their source used for the characterisation of invasive behaviour.

Primarily, I optimised the invasion assay by conducting the assay under a gradient of 0%-10% serum conditions. The gradient provides a chemo-attractant for cells to migrate and invade the Matrigel. This is a common method to study cell invasion, however in my studies this resulted in invasion of the ORI-3 cell line. This cell line, derived from normal epithelium was used as a negative control for invasion. I therefore chose to conduct invasion assays using low serum conditions (1% Foetal calf serum) without the addition of chemoattractants, therefore only invasive cells would invade through the matrix. The invasion assay therefore provides an accurate representation of the invasive potential of the respective cell line. Each assay was conducted in triplicate and on three separate occasions. The chambers were fixed and stained with GIEMSA and the total number of invaded cells counted in 5 different fields of each chamber and the average across the triplicate experiments compiled (Figure 26).

I used low serum conditions and reduced growth factor Matrigel for the invasion assay since growth factors and mitogens can attenuate the assay. In addition, replication allows for increased accuracy and addition of error bars to the graphical results. The results from the invasion assay clearly show that for each cell line analysed, there is an increase in the number of invasive cells after 48 hours. The ORI-3 cell line, which represents normal epithelium, displayed the least degree of invasion, a result which was to be expected. Both the K1 and FTC236 cell lines displayed the greatest degree of invasion after 24 hours, even more so after 48 hours. This was followed by BCPAP and FTC133 which also displayed significant invasion. The Hth74 cell line did not display significant invasion of the matrix. The increase in invasion after 48 hours may be due to cell proliferation and not cell invasion. Taken together, the papillary carcinoma cell line, K1 and the metastasis of

follicular carcinoma, FTC236 displayed the greatest degrees of invasion. This enabled the characterisation of both negative (ORI-3) and positive control (K1 and FTC236) for future experiments.

These results show the invasive potential of thyroid cell lines. In particular, the ORI-3 cell line displayed basal levels of invasion a result which was expected since this cell line is derived from normal epithelium, and therefore is not expected to invade, and also shows the invasion assay is functioning properly. This clearly marks this cell line as the negative control for future experiments. From the carcinoma cell lines, the K1 cell acted as the superior positive control, as it displayed the greatest degree of invasion both after 24 hours and 48 hours.



**Figure 26.** The invasive potential of cell lines derived from different thyroid pathologies using the Boyden chamber system. Invasion assays displaying percentage of invaded cells after 24 hours (blue) and 48 hours (red) respectively. Cells were seeded at a density of  $3 \times 10^5$ /insert. Each cell line was assayed three times, each time in triplicate Boyden chambers. After 24 and 48 hours respectively, invaded cells were visualised using GIEMSA stain, and 5 different fields of view for each were photographed and the average amount of cells counted. **(A)** Actual cell numbers of invaded cells. T-test analysis was conducted by comparing ORI-3 cells with each of the cell lines at 24 and 48 hours respectively. Significant p-values are indicated above the bars together with black stars. **(B)** Cell numbers were converted into percentages and graphically presented. 100% invasion was assumed at the 48 hour time point of the K1 sample.



### **6.3 Characterising the role of Slug and OPN in tumour invasion**

My expression profile of Slug and OPN revealed an increased expression in the RASV12 model. This FA model also displayed prominent invasive capability. This may be linked to the up-regulation of these respective genes. Indeed, the up-regulation of both Slug and OPN has been highlighted in the literature in a number of tumours including PTC.

Literature analysis revealed OPN to be over-expressed in human PTC samples compared with normal epithelium, FA and multinodular goiters. Also, OPN was shown to induce invasiveness in PTC cell lines (Guarino, Faviana et al. 2005). Details regarding molecular mechanisms, processes, phenotypic effects and up-stream regulators and down-stream effectors of OPN require further investigation to further elucidate the role of OPN in thyroid tumourigenesis.

Similarly Slug, a transcriptional repressor has been implicated in thyroid tumourigenesis. Its expression was detailed in a number of thyroid cell lines. Expression of Slug is minimal in the ORI-3 cell line, a cell line which is derived from normal thyroid epithelium, and it is over-expressed in cell lines derived from thyroid carcinomas (Hardy et al. 2007). Morphologically, cells expressing the transcription repressor Slug did not form close contacts and Slug was deemed functionally active based on the localised nuclear expression. Although the work did establish an implication of Slug in thyroid carcinogenesis, there was limited experimental data on the background mechanisms of Slug and its link with carcinogenesis. Slug is a known transcriptional repressor of E-cadherin, and E-cadherin down-regulation is associated with the induction of EMT. This study will reflect on the previous work pertaining to Slug and its role in cell-cell contact, but also its role in cell-substrate

interaction. In addition, the section details the expression of Slug and OPN in thyroid cell lines, their impact on the invasion, with detailed molecular background surrounding the impact.

### *6.3.1 Characterisation of Slug in thyroid cell lines*

To assess if the expression of Slug correlates with the level of invasion, I determined expression of Slug in a range of thyroid cell lines derived from different tumours. Real-time PCR (qPCR) was used to quantify Slug mRNA within each cell line. The method entailed extracting total RNA from sub-confluent cells using Trizol which was subsequently processed by Dr. Claudia Consoli from CBS who performed the real-time reverse transcription of Slug using GAPDH as a loading control. The qPCR analysis revealed ORI-3 cells to have no Slug mRNA. However, K1 cells had ~250 units of normalised expression whilst in the other cell lines the quantity was considerably less at ~50 units of normalised expression (Figure 27A).

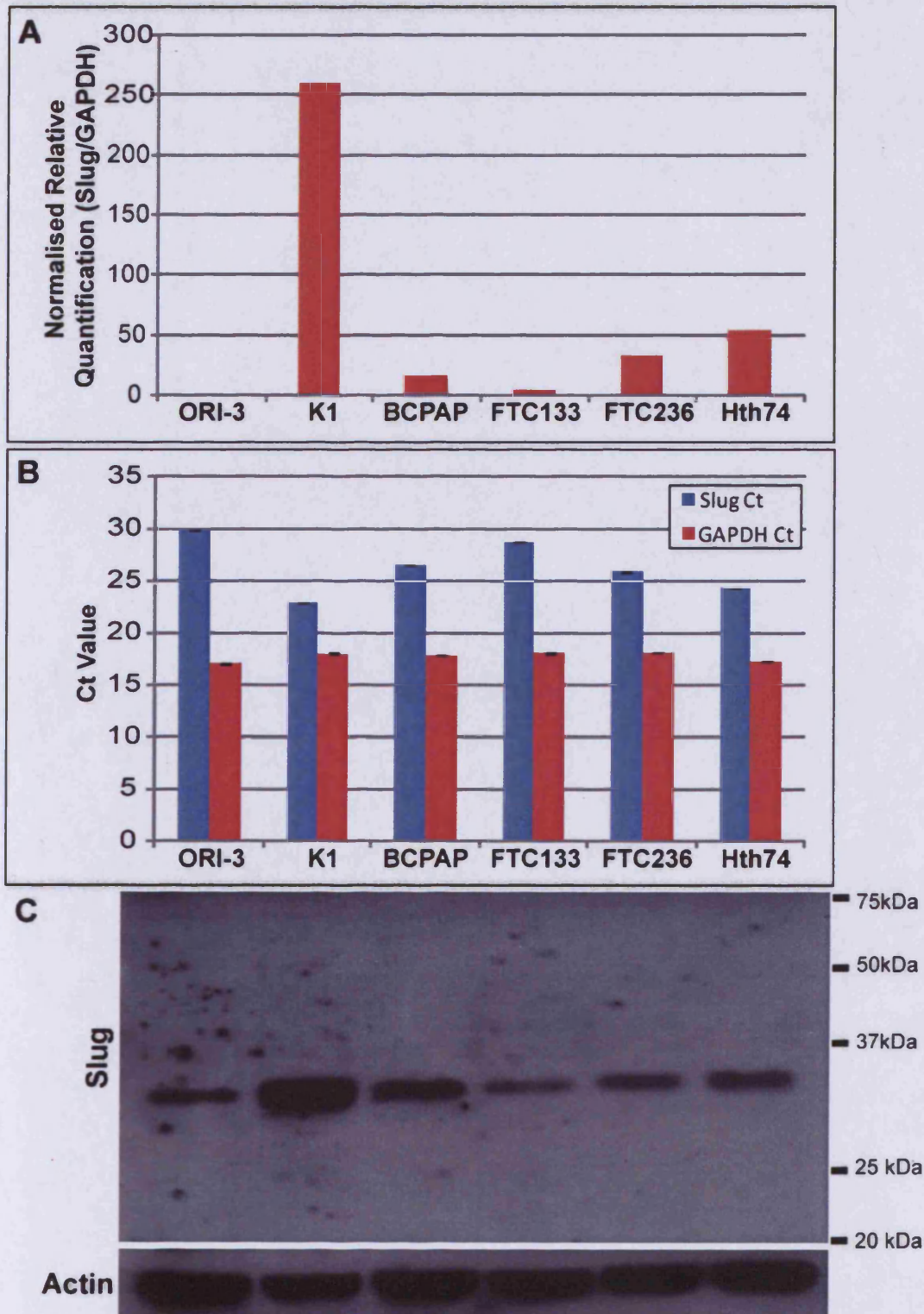
In parallel, I extracted protein lysates from sub-confluent dishes of each of the respective cell line. Samples were quantified and SDS-PAGE western blot analysis was conducted. The lysate samples were ran on a 10% acrylamide gel and transferred to a PVDF membrane. The membrane was then probed with a Slug specific antibody and re-probed with actin, as a loading control for the western blot.

Contrasting with the qPCR analysis, the western blot analysis revealed basal Slug expression in ORI3 cells. The FTC133, FTC236 and Hth74 cell lines also had similar expression levels of Slug. The most protein was detected in the papillary carcinoma

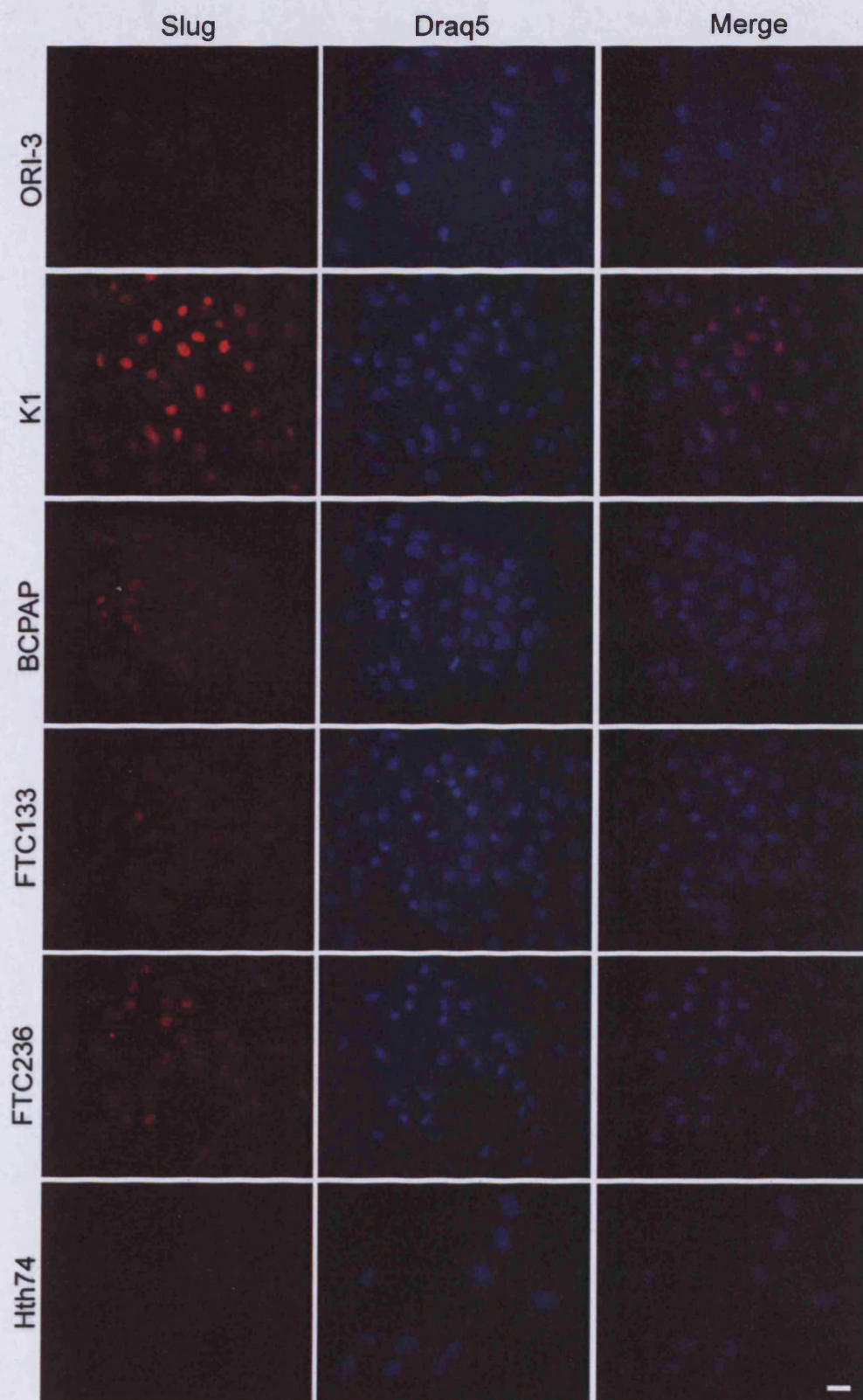
cell lines K1 and BCPAP respectively, with less detected in the BCPAP cell line (Figure 27C).

To observe cellular localisation of Slug and a qualitative measure of Slug expression immunofluorescence was utilised. For immunofluorescence, cells were grown onto glass coverslips, fixed in 4% paraformaldehyde, permeabilised with 0.1% Triton-X and probed with a Slug specific antibody. Nuclei were counter-stained using the DRAQ5 stain.

Immunofluorescence analysis revealed the Slug to be localised at the nucleus, a result consistent with previous studies (Hardy et al. 2007). In addition the signal intensity of Slug can be used a qualitative measure of Slug expression. The K1 cell line had the greatest signal intensity which indicates an increased expression of Slug in this cell line (Figure 28). This work details the expression of Slug in a range of thyroid cell lines derived from different pathologies. Slug was detected at basal levels in ORI-3 cells, whilst in the PTC cell line K1, Slug was detected at significantly higher levels.



**Figure 27.** Characterisation of Slug in thyroid cell lines **A**, qPCR of Slug in thyroid cell lines normalised to GAPDH. Total RNA was extracted from each cell line and quantified in triplicate. **B**, Ct values of Slug (blue bars) and GAPDH (red bars) obtained from the qPCR analysis. **C**, Western Blot analysis of Slug in each cell line using a Slug specific antibody. Molecular weight markers are highlighted in red. The membrane was stripped and re-probed with actin to observe loading.



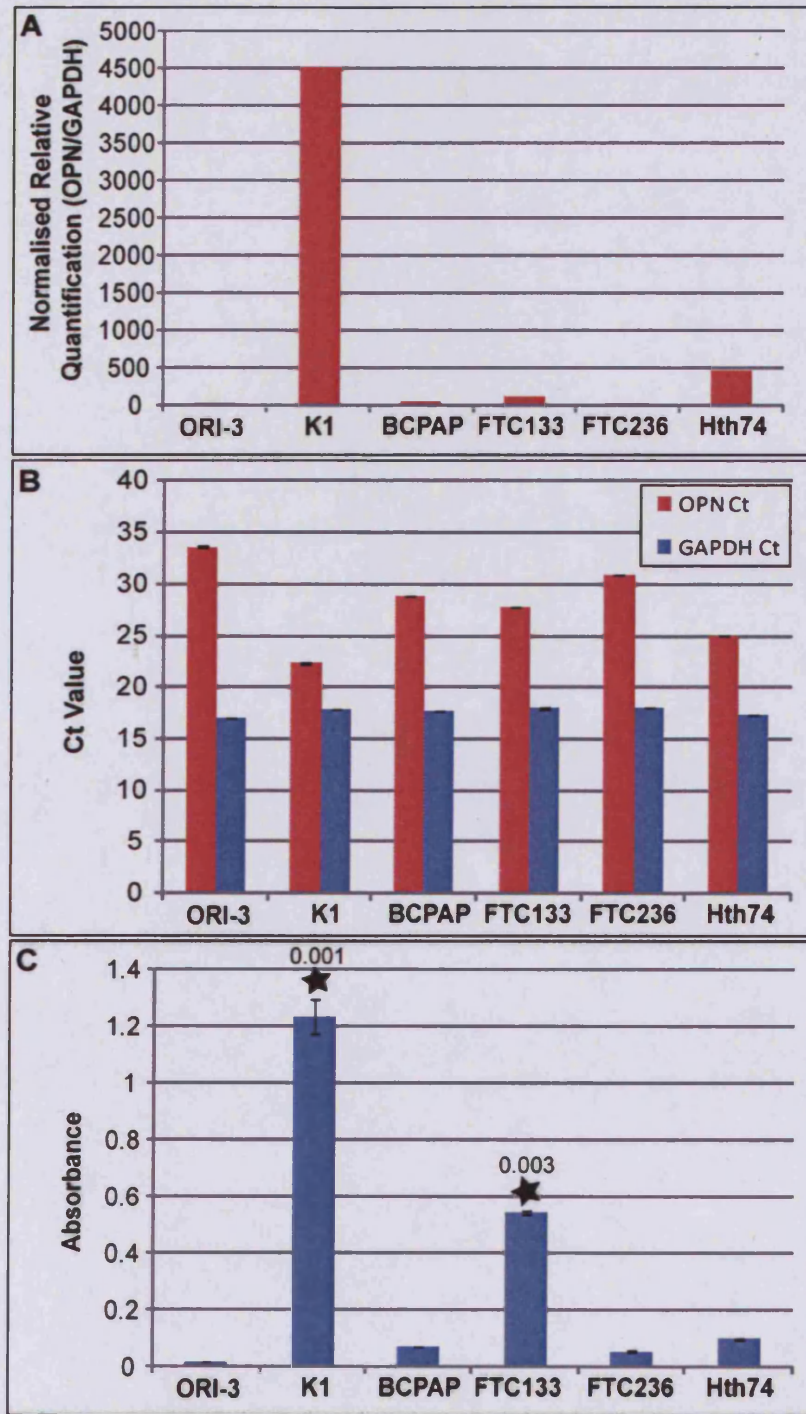
**Figure 28.** Immunofluorescent analysis of Slug expression in cell lines derived from various thyroid pathologies. Cells were stained with Slug (red) and counterstained with DRAQ5 nuclear stain (blue). Scale bar indicates 50 $\mu$ m and is applicable to all images.

### 6.3.2 Characterisation of OPN in thyroid cell lines

To assess a possible role of OPN with invasion, OPN expression was assessed in the same panel of cell lines used to assess Slug expression. To quantify OPN mRNA content, qPCR was used. I extracted total RNA using Trizol from matched samples that were used for the ELISA analysis. The qPCR showed a massive increase in OPN mRNA in the K1 cells compared with the other cell lines. In ORI-3 cells, OPN was undetectable indicating minimal expression of the OPN protein. In the K1 cells however, OPN was highly expressed. The other cell lines had basal level expression of OPN (Figure 29A).

An OPN ELISA was used to detect OPN protein expression. Cells were grown to 80% confluence and the conditioned media was collected from the cells after 48 hours and analysed by a commercially available Enzyme-linked immunosorbent assay (ELISA) assay directed towards OPN. The ELISA was conducted in triplicate and the absorbance's averaged. The results correlated with the mRNA content with K1 displaying the highest absorbance. This target could potentially regulate invasion of K1 cells and further experiments were conducted to elucidate this theory (Figure 29C).





**Figure 29.** Characterisation of Osteopontin in thyroid cell lines. **A**, qPCR was conducted on thyroid cell lines using RNA extracted from each cell line. Quantification was conducted in triplicate from a single sample normalised to GAPDH. **B**, Ct values of OPN (blue bars) and GAPDH (red bars) obtained from the qPCR analysis. **C**, ELISA was conducted on matched samples using a commercially available assay system. In triplicate, conditioned media was collected from matched samples and analysed with the ELISA assay system. T-test analysis was conducted by comparing ORI-3 cells with each of the cell lines at 24 and 48 hours respectively. Significant p-values are indicated above the bars together with black stars.

#### **6.4 The *shRNA mediated knockdown of Slug and OPN***

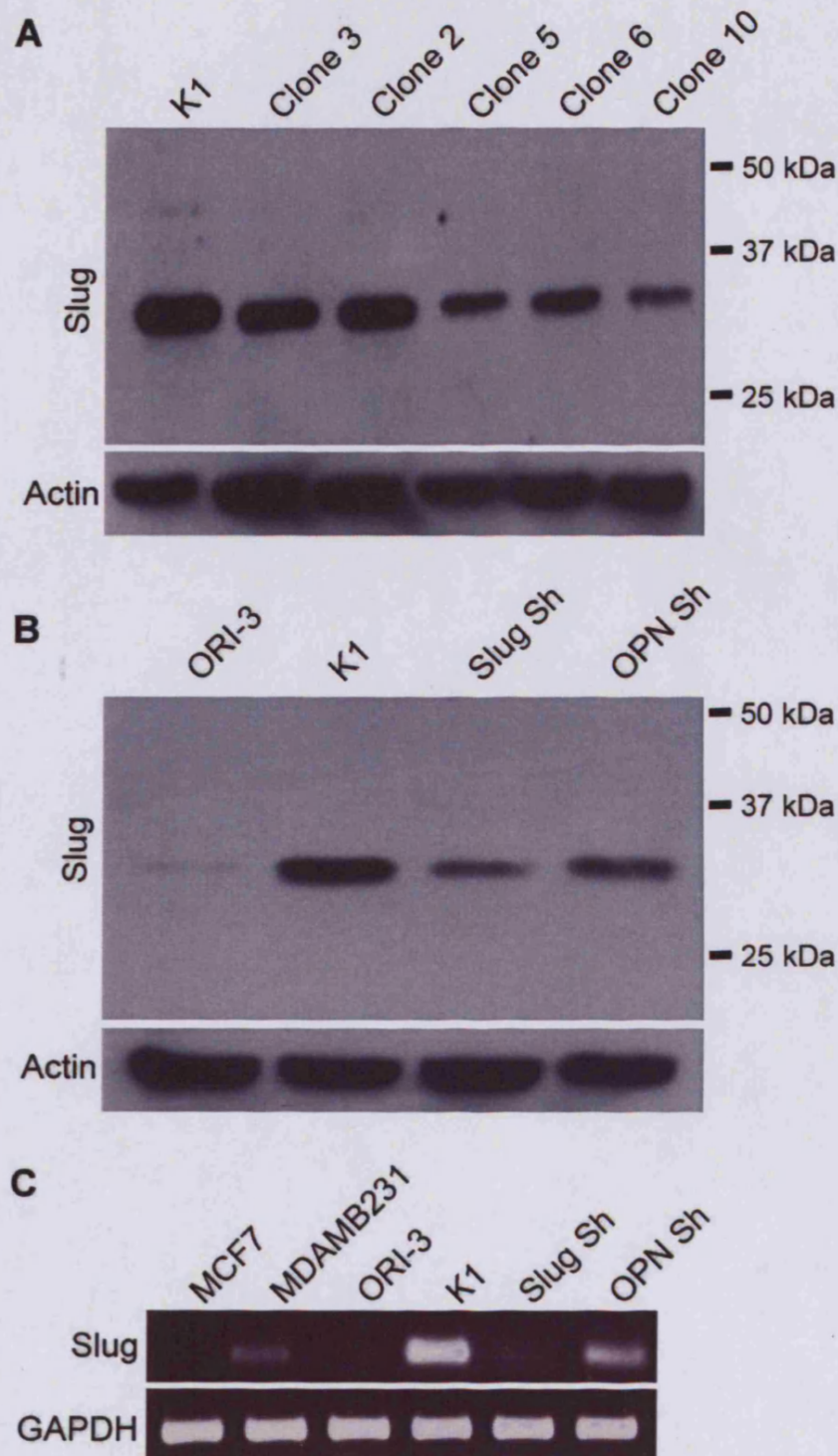
Since I had successfully highlighted a link between Slug and OPN expression with invasive capability I next sought to investigate if this was the case. To achieve this aim I genetically manipulated Slug and OPN expression in the K1 cell line by conducting knock-down experiments.

Plasmid glycerol stock targeting Slug was obtained from Sigma-Aldrich from their MISSION™ shRNA archive. The 293FT cell line was used to generate high-titer Slug lentivirus with the use of the ViraPower™ Lentiviral Expression System. Subconfluent dishes of K1 cells were infected with lentiviral particles expressing shRNA directed towards Slug and OPN respectively. Cells infected with the shRNA were puromycin resistant.

Using the K1 lysate as a positive control, western blot analysis detected Slug expression in the clones expressing the Slug shRNA. This confirmed the most prevalent knockdown of Slug in clone 10 (Figure 30A). Further confirmation of the knockdown was conducted using protein lysates, along with ORI-3 (negative control) and K1 (positive control). This re-confirmed the knockdown of Slug expression in clone 10 (Figure 30B). This clone was selected for future experiments and so was named Slug Sh.

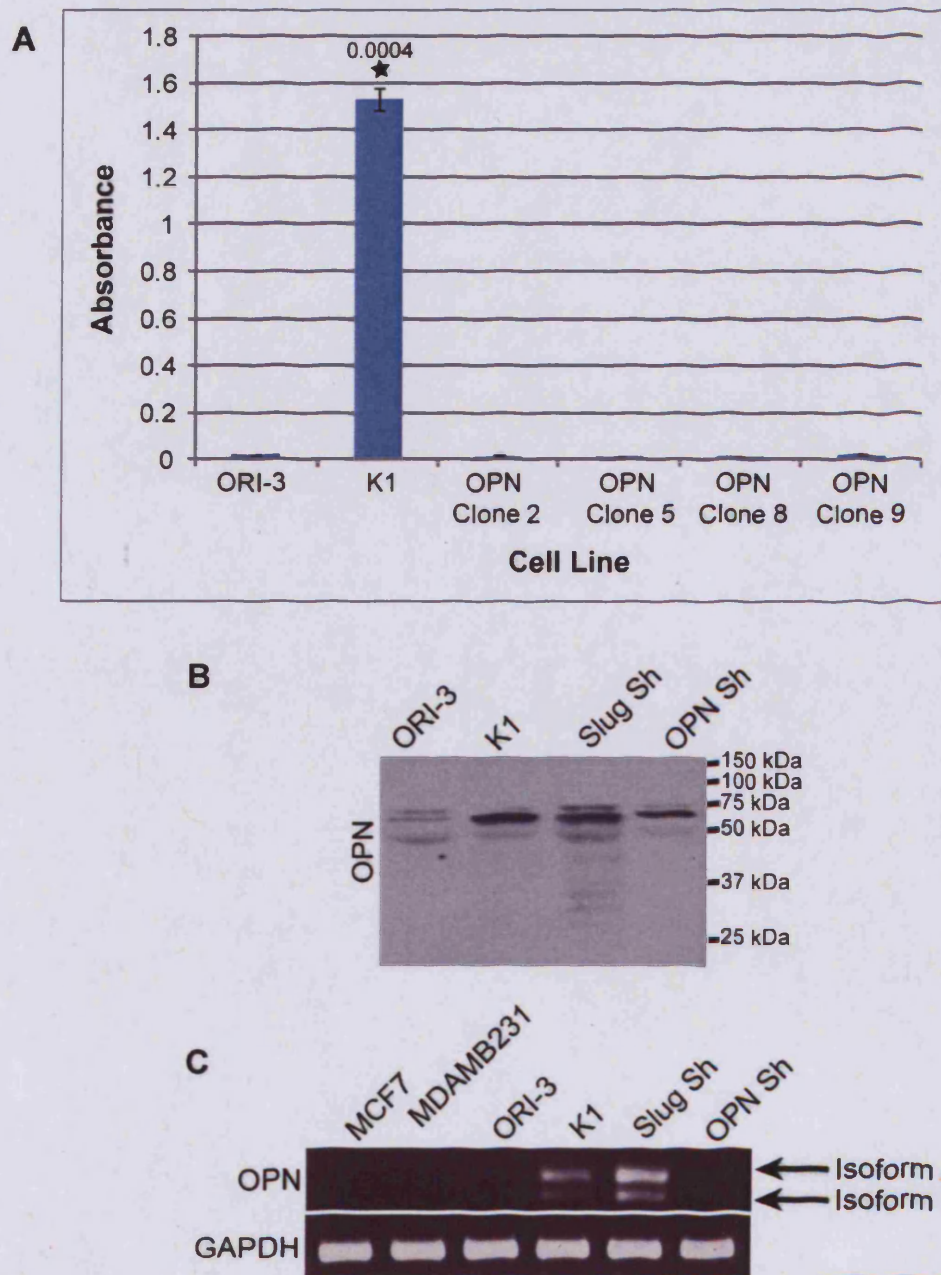
Observing mRNA content reconfirmed the western result of Slug knockdown in K1 cells (Figure 30C).





**Figure 30.** Confirmation of shRNA mediated Slug knockdown. **A**, Confirmation of shRNA knockdown of Slug. Western blot of K1 (positive control) and Slug knockdown clones observing Slug expression. **B**, Western blot analysis re-confirming Slug knockdown in K1 cells with expression of Slug in OPN Sh. ORI-3 (negative control) and K1 (positive control) to reconfirm knockdown of Slug in K1 cells with actin as a loading control. **C**, RT-PCR was conducted using Slug specific primers and reverse transcriptase to semi-quantify mRNA content of Slug in cell lines. MCF7 and MDAMB231 were used as additional controls.

In parallel to the shRNA knockdown of Slug, OPN expression was also knocked down using the same method previously detailed. The system of commercially available glycerol stock of shRNA directed towards OPN was obtained and used. The ViraPower™ Lentiviral Expression System with the 293FT cell line was used to package the lentiviral shRNA into particles. The OPN protein is typically secreted, therefore conditioned media was obtained from OPN clones. Briefly, the OPN clones were grown to sub-confluence and the conditioned media collected and analysed by ELISA. The results show as previously a high level of OPN in K1 cells with basal level of OPN in ORI3 cells. All OPN ShRNA clones showed a dramatic reduction of OPN, highlighting efficient knockdown of OPN. Any of the clones could be suitable for future experiments; however clone 2 was selected (Figure 31A), and was referred to as OPN Sh for future experiments. Using conditioned media with an OPN specific antibody, western blot analysis revealed a decrease in OPN in clone 2 (Figure 31B). Interestingly, the detection of OPN with western blot and ELISA did not correlate with each other, but this may be due to the sensitivity of the assays or experimental procedures. Since conditioned media was utilised in the western blot, an appropriate loading control could not be gained. In ORI-3, MCF7 and MDAMB231 cells there was basal level expression (Figure 31C). Observing OPN mRNA content, there are clearly 2 dominant OPN isoforms identified. The shRNA targets all the OPN isoforms, and this was confirmed as OPN mRNA was not detected in OPN Sh. Consistent with both the ELISA and western blot analysis ORI-3 cells displayed little or no OPN. The RT-PCR analysis of OPN detected isoforms A and B of OPN, but this detection method failed to detect isoform C. The mRNA content in OPN Sh cells showed no detectable OPN isoforms (Figure 31C). OPN mRNA was only detectable in K1 cells and Slug Sh cells.



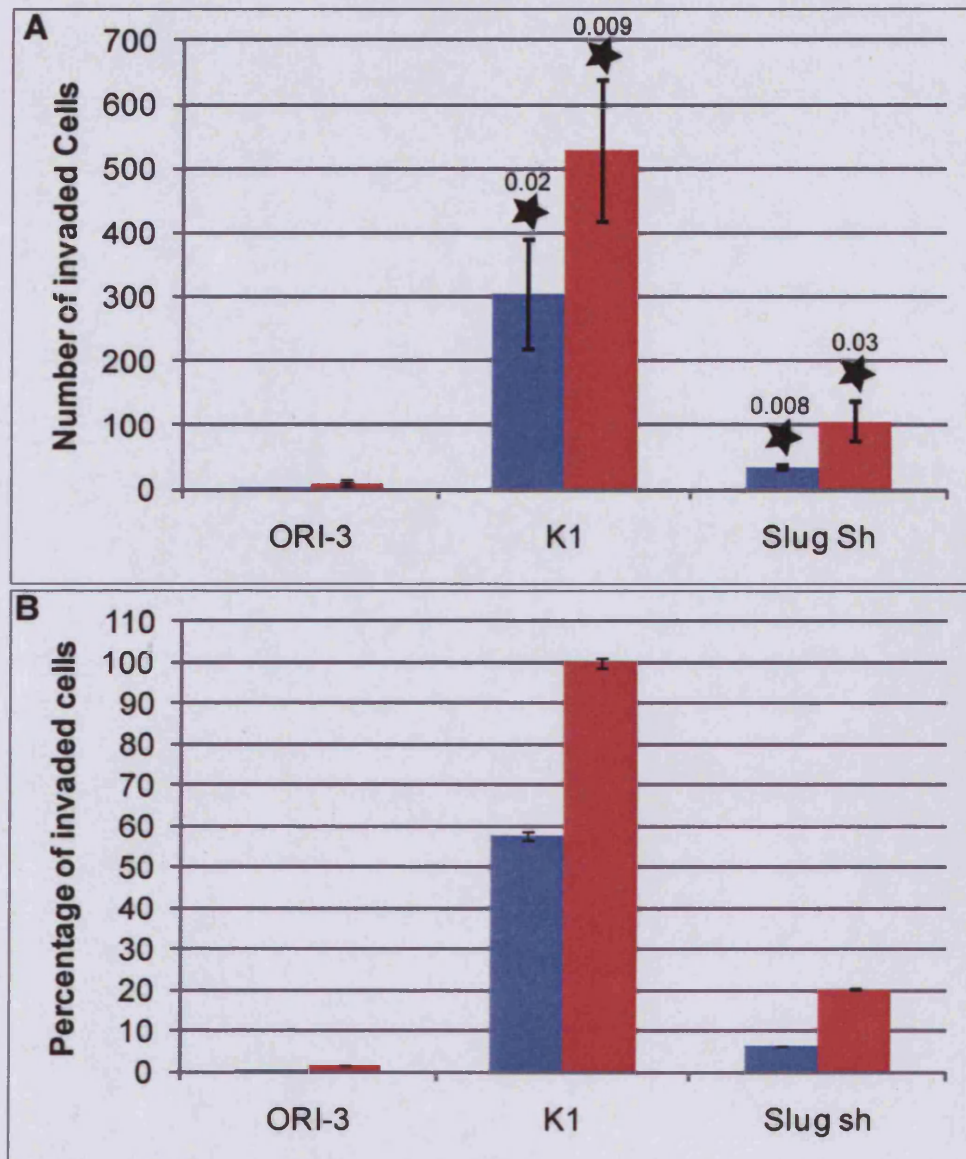
**Figure 31.** Confirmation of shRNA mediated knockdown of OPN. **A**, ELISA was conducted using a commercially available assay system. Conditioned media was collected from clones, with K1 and analysed with the ELISA assay system. Each sample was analysed in triplicate. ORI-3 is a negative control, K1 acts as a positive control and remaining samples depict cells with knockdown of Osteopontin. T-test analysis was conducted by comparing ORI-3 cells with each sample. Only the K1 cell line displayed significance ( $p=0.004$ ) indicated above the bar together with a black star. **B**, Western blot using conditioned media loaded equally. **C**, RT-PCR was conducted using OPN primers which are able to detect all isoforms of OPN. GAPDH acts as a loading control for the PCR experiment. MCF7 and MDAMB231 were used as additional controls. ORI-3 was the thyroid negative control and K1 the thyroid positive control.



I next aimed to determine if the knockdown of Slug or OPN affected the invasive capability of the K1 cell line. I used the same *in vitro* invasion assay system with the same experimental conditions. The assays like previous were scored at 24 hours and 48 hours post invasion, and the results summarised as number of invaded cells and percentage invasion.

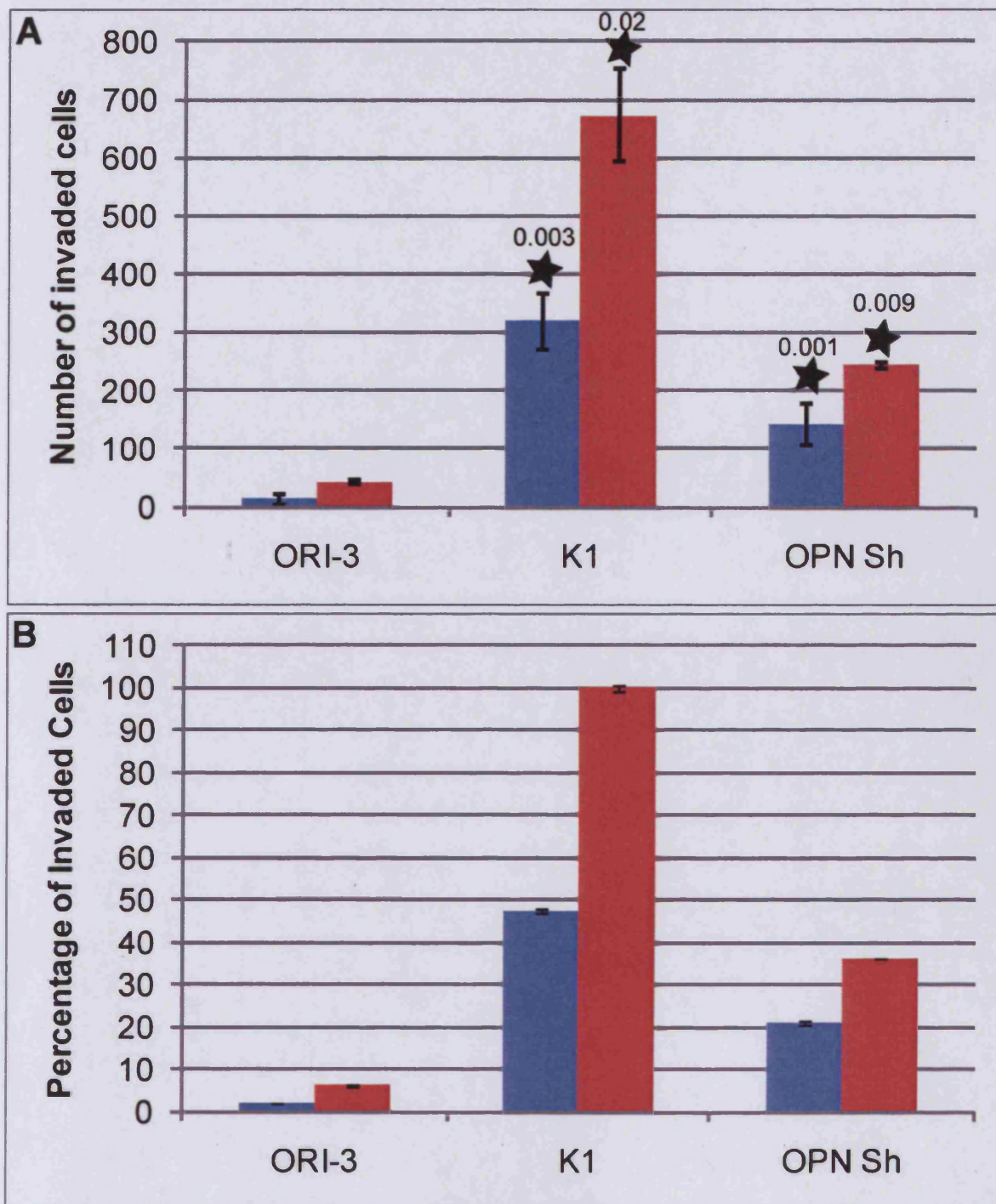
The K1 cell line at the 48 hour time point displayed the highest degree of invasive behaviour and was taken as 100% invasion. The Slug Sh and OPN Sh timepoints were calculated as the percentage of invasion by comparing with K1 at 48 hours.

After a period of 48 hours in ORI-3 cells, the invasion assays showed minimal invasion with less than 2% of invaded cells detected. In K1 cells, there was prominent invasion after 24 hours with 57% invaded cells, with maximum invasion after 48 hours. With the Slug Sh sample, there was a dramatic decrease in invasion both after 24 hours and 48 hours with 6.5% invaded cells detected after 24 hours with an increase to 20% cells after 48 hours. This result provides a link between Slug expression and the invasive behaviour of cells (Figure 32).



**Figure 32.** The invasive potential of the Slug Sh cell line. Invasion assays displaying **A)** average total invaded cell number and **B)** percentage of invaded cells. Invasion assays were taken after 24 hours (blue) and 48 hours (red) respectively. Each cell line was assayed three times in triplicate Boyden chambers. Invaded cells were visualised using GIEMSA stain, and 5 different fields of view for each were photographed and the average amount of cells counted. Values were then converted into percentages and graphically presented. At the 48 hour time point, 100% total invasion in the K1 cell was assumed. The other samples were calculated for percentage invasion from the total invasion of K1. T-test analysis was conducted comparing ORI-3 with either K1 or Slug Sh at 24 hours and 48 hours respectively. Significance is shown as black stars above the bar with the *p*-value.

To establish if OPN expression was linked to invasive behaviour, invasion assays were conducted. After a period of 24 hours ORI-3 cells like previous results displayed minimal invasion with 6.2% invaded cells detected after 48 hours. In K1 cells, there was prominent invasion after 24 hours ( $p=0.003$ ) with 47% invaded cells with maximum invasion after 48 hours ( $p=0.02$ ). The knockdown of OPN had a dramatic effect on invasive behaviour as 21% invaded cells were detected after 24 hours ( $p=0.001$ ) with 36% after 48 hours ( $p=0.009$ ). This experimental data, like the data of Slug Sh provided an interesting link of OPN regulating the invasive behaviour of the papillary carcinoma cell line, K1. (Figure 33).



**Figure 33.** The invasive potential of the OPN Sh cell line. Invasion assays displaying **A)** average total invaded cell number and **B)** percentage of invaded cells. Invasion assays were taken after 24 hours (blue) and 48 hours (red) respectively. Each cell line was assayed three times in triplicate Boyden chambers. Invaded cells were visualised using GIEMSA stain, and 5 different fields of view for each were photographed and the average amount of cells counted. Values were then converted into percentages and graphically presented. At the 48 hour time point, 100% total invasion in the K1 cell was assumed. The other samples were calculated for percentage invasion from the total invasion of K1. T-test analysis was conducted comparing ORI-3 with either K1 or OPN Sh at 24 Hours and 48 hours respectively. Significance is shown as black stars above the bar with the *p*-value.



## 6.5 Summary

To study tumour invasion cell lines derived from various thyroid cancers were used due to their unlimited proliferative potential and their representation of tumour pathology. This work was initially carried out in cell lines with a view to conduct experiments on primary models. An *in vitro* invasion assay system was utilised to study the concept of tumour invasion. High serum conditions attenuate the assay and induces invasion in non invasive cell lines including ORI-3 and MCF10A (data not shown). Therefore, invasion assays were conducted using growth factor reduced Matrigel with low serum culture conditions to limit the effect of growth factors and stimulants. The reconstituted basement membrane matrix offered a consistent representation of the basement membrane for the analysis of tumour invasion.

The thyroid cell lines analysed included the PTC cell lines K1 and BCPAP, which carry the BRAF<sup>V600E</sup> mutation. This mutation is associated with increased aggressiveness (Kebebew, Weng et al. 2007). The invasive capability of the K1 cell line correlates with the aggressiveness of *in vivo* tumours. ORI-3 cells are derived from normal epithelium and do not show an invasive capability, which is consistent with primary tumours. The FTC133 cell line derived from FTC, displayed increased invasion, but its metastasis FTC236 displayed a significant increase in invasive potential. This would be expected as metastatic cells would phenotypically have a higher potential to invade the basement membrane. This also highlights the invasion assay was functioning as it should with degrees of invasion displayed by the cell lines analysed. The anaplastic cell line, Hth74 displayed a low degree of invasion implying an anomalous result, as anaplastic carcinomas are highly invasive and aggressive and induce fatality within a year (Sherman 2003). The Hth74 cell line may



display a high degree of invasive behavior and may have invaded through the Matrigel before the 24 hour time point.

I next investigated if either Slug or OPN were involved in regulating an invasive phenotype. To achieve this aim, the expression of Slug and OPN was characterised across the same panel of cell lines. Both genes had increased expression in the K1 PTC cell line, recapitulating evidence presented in other studies. Literature analysis has previously eluded the potential role of these two genes in PTC. Over expression of OPN was demonstrated in primary PTC samples and was involved with the invasive phenotype of PTC (Guarino, Faviana et al. 2005). Slug has also been highlighted in PTC cell lines (Hardy, Vicente-Duenas et al. 2007).

Taken together, the K1 cell line displayed the highest expression of both Slug and OPN together with the highest degree of invasive behavior. The ORI-3 cell line which displayed minimal invasion displayed basal level of expression of Slug and OPN. This indicates prominent expression of both Slug and OPN may be linked to increased invasive ability.

To investigate the role of Slug and OPN in tumour invasion, I chose to genetically manipulate gene expression by knocking down expression in the K1 cell line. To conduct shRNA-mediated knockdown of I purchased glycerol stocks of shRNA and prepared the lentiviral particles to transduce into the K1 cell line. Lentiviral vectors were chosen as lentiviruses can target both dividing and non-dividing cells. To determine effective knock-down, expression of Slug and OPN was re-characterised allowing me to select clones displaying reduced expression of Slug and OPN. These clones were termed Slug Sh and OPN Sh respectively.

Observing OPN mRNA in OPN Sh by RT-PCR revealed the identification of multiple isoforms of OPN. In a separate study of mesothelioma, OPN isoforms were

assessed for their tumourigenic properties. The OPN gene produces three RNA isoforms which encode three protein isoforms. Isoform-A is the full length isoform of OPN, isoform-B excluding exon 4 and isoform-C which excludes 3. In mesothelioma, isoforms-A and -B were up regulated, whilst isoform C was not. Isoforms-A and -B induced tumourigenic phenotypes including proliferation, migration, invasion and colony formation (Ivanov, Ivanova et al. 2009). Isoform-C was not detected in the RT-PCR analysis, but considering the role of OPN in mesothelioma this isoform can be deemed insignificant in terms of its tumourigenic activity.

Interestingly in K1 cells, isoform-A was more abundant in K1 cells, and following lenti-viral knockdown of OPN, no isoform of OPN was detected by RT-PCR with reduced expression detected by western blot. In Slug Sh, the mRNA content of OPN was increased. Currently there is no documented link between Slug and OPN. The down-regulation of OPN mRNA in Slug Sh and up-regulated Slug mRNA in OPN Sh indicates a possible interaction of Slug and OPN. The results showed an increase of mRNA content in Slug Sh compared to K1 cells. This could be due to slight unequal loading during the RT-PCR analysis, or it may be caused by a direct regulatory link between Slug and OPN which results in the increased expression of OPN. To provide evidence for a direct link between Slug and OPN, OPN Sh cells were analysed for Slug expression.

To study the effect of reducing expression of Slug and OPN on invasive behavior, I conducted invasion assays using K1 as a positive control and ORI-3 cells as the negative control. Overall, the invasive behavior was significantly reduced for both OPN Sh and Slug Sh. This strongly suggests the expression of Slug and OPN is linked to invasive behavior. The K1 cell line displayed increased invasive behavior

with the ORI-3 cell line displaying minimal invasive behavior. This recapitulates the data which was obtained earlier.

## **6.6 Invasive processes mediated by Slug and OPN**

After establishing a role for Slug and OPN in thyroid tumour invasion, I sought to investigate the cellular mechanisms mediated by Slug and OPN which induce invasion.

### **6.6.1 Proteinase activity mediated by Slug and OPN**

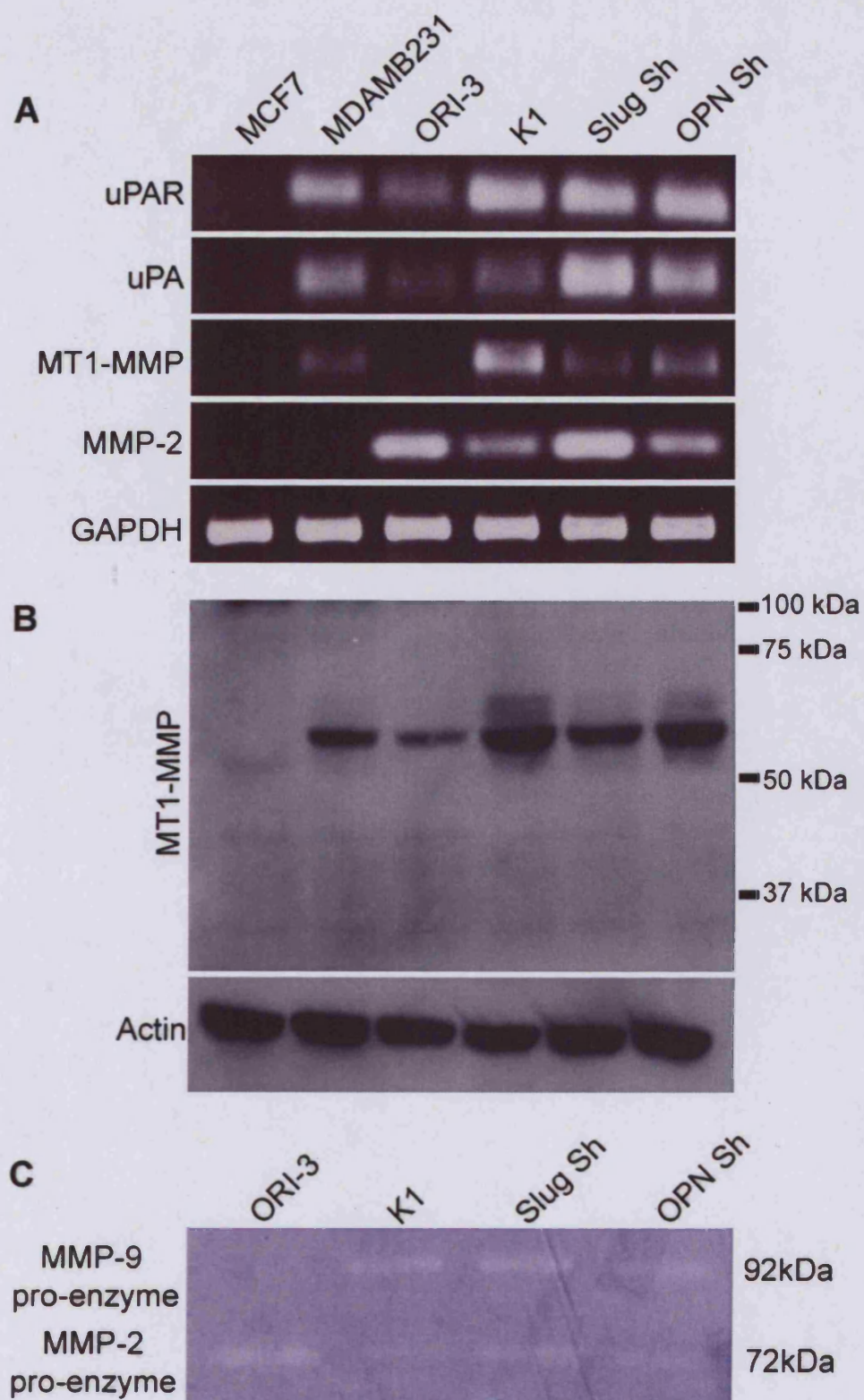
Proteinase activity has been highly regarded as mechanism which aids invasive behaviour of tumours. To investigate the effect of proteinase activity in thyroid tumour invasion and the relation to the expression of Slug and OPN, I conducted RT-PCR analysis using specific primers targeting the urokinase plasminogen activator, its receptor, MMP-2 and MT1-MMP. In parallel, MT1-MMP protein expression was analysed by western blot. To assess MMP activity, gelatin zymogram analysis was conducted to observe MMP-2 and MMP-9 pro- and active forms.

The ORI-3 cell line displayed basal level detection of uPA and uPAR. MT1-MMP was not detected, whilst MMP-2 was. Since ORI-3 does not display an invasive capability, MMP-2 mRNA alone is not sufficient to induce invasion. In the K1 cell line, MT1-MMP mRNA content was the highest compared to the other cell lines. The uPA receptor was detected considerably more in K1 cells compared with ORI-3 cells. This may indicate a potential role of the uPAR in the invasive capability of K1 cells, although this would require further investigation to ascertain if this is the case. The Slug Sh and OPN Sh cells displayed a comparable mRNA content of uPAR to the K1 cell line. The mRNA content of uPA was detected at basal levels in both ORI-3 and K1 cells, indicating the plasminogen activator system is not applicable in these cells.

In Slug Sh and OPN Sh the content of uPA was increased compared to K1 cells. Since uPA requires the binding to uPAR for plasminogen activation, it implies the system does not contribute to the invasive potential, and indicates that in PTC the urokinase plasminogen activation system does not have a role in invasion (Figure 34A).

Western blot analysis of MT1-MMP expression correlated with mRNA content. The K1 cell line displayed the highest degree of expression, with ORI-3 and MCF7 the lowest. MT1-MMP expression in the Slug Sh cells was apparently reduced. This indicates a potential regulatory mechanism involving MT1-MMP and Slug. MT1-MMP protein expression in OPN Sh cells was not significantly altered, although the mRNA content of MT1-MMP was reduced (Figure 34B).

The Gelatin zymogram analysis was conducted to observe MMP-2 and MMP-9 activity using conditioned media collected from samples which were used for the western blot analysis. To conduct this assay the zymogram was loaded with conditioned media, electrophoresed and stained using coomassie blue. Based upon size, the digested areas in the gel represent latent and active forms of MMP-2 and MMP-9. For my study, only MMP-2 and MMP-9 pro-enzymes were detected in the cell lines, indicating MMP-2 and MMP-9 activity does not contribute to the invasive process in Slug or OPN mediated invasion (Figure 34C).



**Figure 34.** Analysis of proteinase activity in Slug Sh and OPN Sh. **A.** RT-PCR using specific primers directed to uPAR, uPA, MT1-MMP and MMP-2 with GAPDH as a loading control. **B.** Western blot of MT1-MMP using cell lysates with actin as the loading control. **C.** Zymogram analysis using conditioned media from matched protein samples detecting MMP-2 and MMP-9 pro-enzymes only.

### 6.6.1.1 Summary

Matrix metalloproteinases (MMPs) are a family of zinc-binding end peptidases which are known to degrade the majority of components of the ECM, which their expression prevalent in malignant tumours. Membrane bound MMPs have a C-terminus hydrophobic domain which anchors the MMP to the cell membrane, thereby restricting its activity to the cell surface. MT1-MMP acts as a master activator of MMP-2, an MMP which degrades type IV collagen in the basement membrane (Sato, Takino et al. 1994). In addition, MT1-MMP expression has been shown to correlate with tumour growth and angiogenesis via the up-regulation of VEGF in adenocarcinoma (Sounni, Devy et al. 2002).

Expression of MT1-MMP was reduced in both the Slug Sh and OPN Sh cells, confirmed by RT-PCR and western blot. This indicates that OPN and Slug may possibly have a role in the regulation of MT1-MMP. In mesothelioma, expression of Slug correlates with the levels of MT1-MMP, MMP-2 and MMP-9 (Sivertsen, Hadar et al. 2006). Additionally in lung cancer, Slug mediated an increase in invasion, suppression of E-cadherin and regulation of MMP-2 expression and activity. The authors failed to analyse MT1-MMP expression (Shih, Tsai et al. 2005). For MMP-2 to become active, an MT1-MMP dimer is required. In my investigation, Slug Sh showed a reduction of MT1-MMP. Preliminary zymogram analysis did not reveal active forms of MMP-2 or MMP-9 in K1 cells, indicating that these cell lines do not possess an active MMP-2 or MMP-9 enzyme. The lack of these enzymes in K1 indicates that MMP-2 and MMP-9 do not have a role in invasion. However, this assay was only conducted once and so a repeat experiment is required. In addition, the experimental conditions may require further optimisation to observe active MMP-2 and MMP-9 enzymes.

The evidence presented indicates the reduction in MT1-MMP expression in Slug Sh cells which may imply a regulatory mechanism involving Slug and MT1-MMP, and this may in part explain why Slug Sh cells have a reduced invasive capacity.

In mammary cancer cells, OPN is able to confer an invasive phenotype by up-regulating MMP-2 and uPA (Mi, Guo et al. 2006). In addition, in murine melanoma cells, OPN induced MT1-MMP expression and activation of pro-MMP-2 (Philip, Bulbule et al. 2001). MT1-MMP down-regulation in OPN Sh confirmed a role of OPN in the induction of MT1-MMP expression.

Taken together it appears that MT1-MMP is directly regulated by Slug and OPN. In migrating cells MT1-MMP is re-localised to the lamellipodia and invadopodia (Nakahara, Howard et al. 1997; Mori, Tomari et al. 2002). This displays the dynamic regulation of MT1-MMP which can occur during cell invasion and migration. It is unclear if this occurs in K1 cells, and requires further investigation. MT1-MMP is able to interact with a wide range of proteins including ECM components, CD44 and  $\alpha v$  integrins (Gingras and Beliveau. 2010). Since CD44 and  $\alpha v$  integrins are ligands for OPN, the downregulation of MT1-MMP in OPN Sh may influence the binding of CD44 and  $\alpha v$  integrins to OPN. For OPN-mediated invasion, CD44 and/or  $\alpha v$  integrin interaction with MT1-MMP may be a pre-requisite for OPN mediated-cell invasion. The down-regulation of MT1-MMP in both Slug Sh and OPN Sh may be influence the production of VEGF. MT1-MMP is known to up-regulate VEGF production (Sounni, Devy et al. 2002). VEGF is a pro-angiogenic factor when upregulation is associated with enhanced tumour growth, and invasion in thyroid cancer (Lin and Chao 2005).

The urokinase activator system has a prominent role in tumour progression. The mRNA content was unaltered in Slug Sh and OPN Sh, which indicates that both



OPN and Slug do not regulate the uPAR mRNA levels. Reduced uPAR would result in a decrease of uPA activation, since the binding of uPA to the receptor is required. The mRNA content of uPA was increased in Slug Sh and OPN Sh compared with K1 cells. As Slug Sh and OPN Sh displayed a reduction in invasive behaviour, the increase in uPA mRNA does not have impact on invasion.

## **6.6.2 Cell migration**

### **6.6.2.1 Analysis of the actin cytoskeleton during cell migration**

Cell motility is a complex and an integral process for invasion requiring the cell to regulate cell-cell adhesion, cell-matrix adhesion and regulation of the cell cytoskeleton. Forward motility requires the formation of actin membrane protrusions coupled with regulation of cell adhesion and retraction from the rear of the cell. Cells move by rearranging the actin cytoskeleton with regulation of cell-matrix interactions to mediate the processes of cell migration and motility.

To investigate cell motility and the role of Slug and OPN in mediating this process I conducted an immunofluorescence time course experiment. This involved growing cells to confluence on coverslips and creating a 'wound' in the monolayer of cells. At each indicated timepoint the cells were fixed and stained for F-actin and DAPI by immunofluorescence. This experiment effectively observed wound closure and the distribution of filamentous actin in cells at the leading edge of the wound.

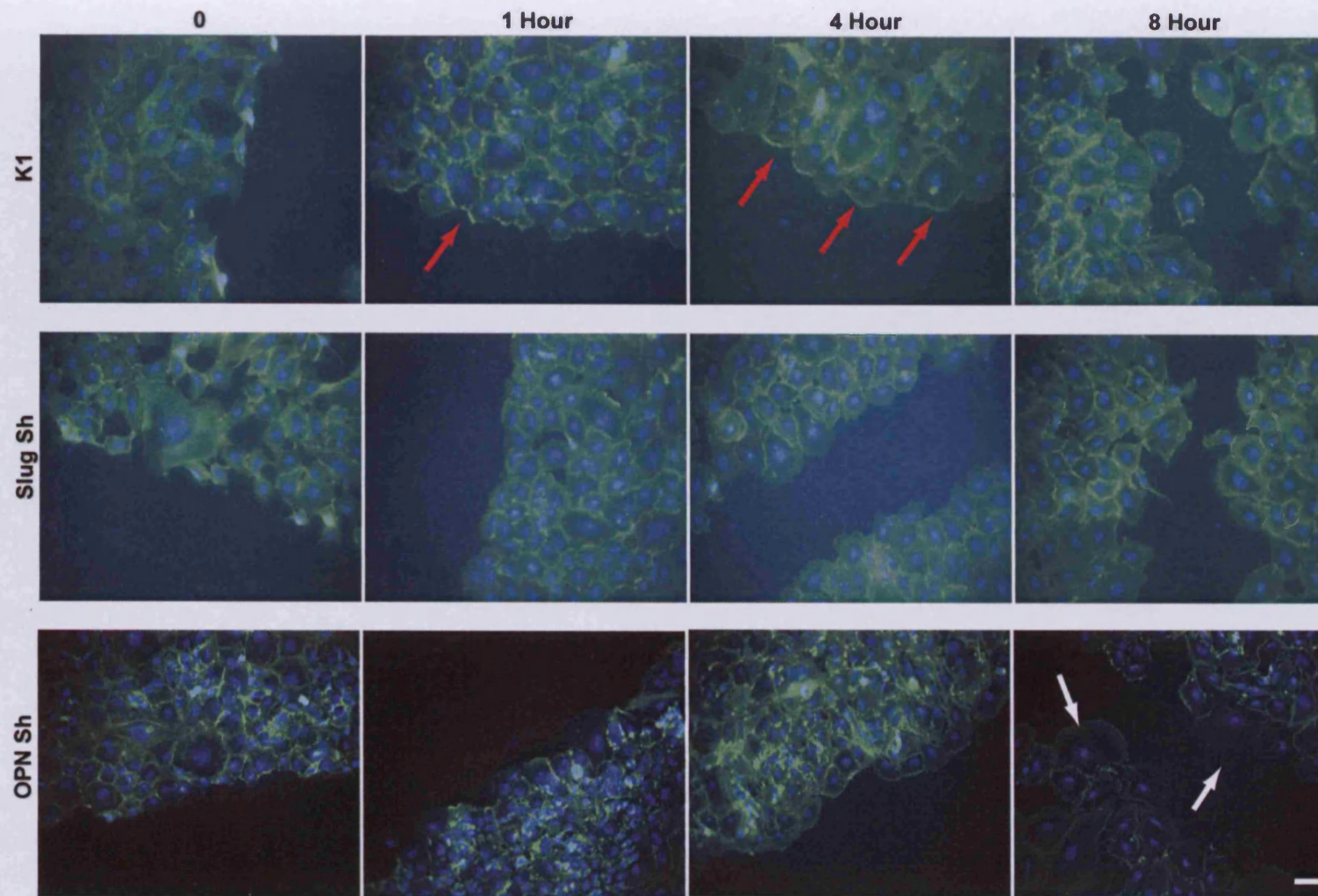
The K1 cells displayed accelerated wound closure at ~12 hours. Both Slug Sh and OPN Sh closed the wound at ~24 hours. At the 4 hour time point, K1 cells have a concentration of F-actin at the leading edge of the wound (Figure 35, red arrows).

Observing actin at each time point revealed the cell size of OPN Sh at the leading edge appeared more expanded compared to K1. For all the samples, actin appeared to be concentrated at the cell membrane.

The cells at the leading edge of OPN Sh displayed an interesting result in that the staining for actin was decreased, highlighting a lack of F-actin stress fibres (Figure 35, white arrows). This is an indication of actin cytoskeleton impairment. In addition,

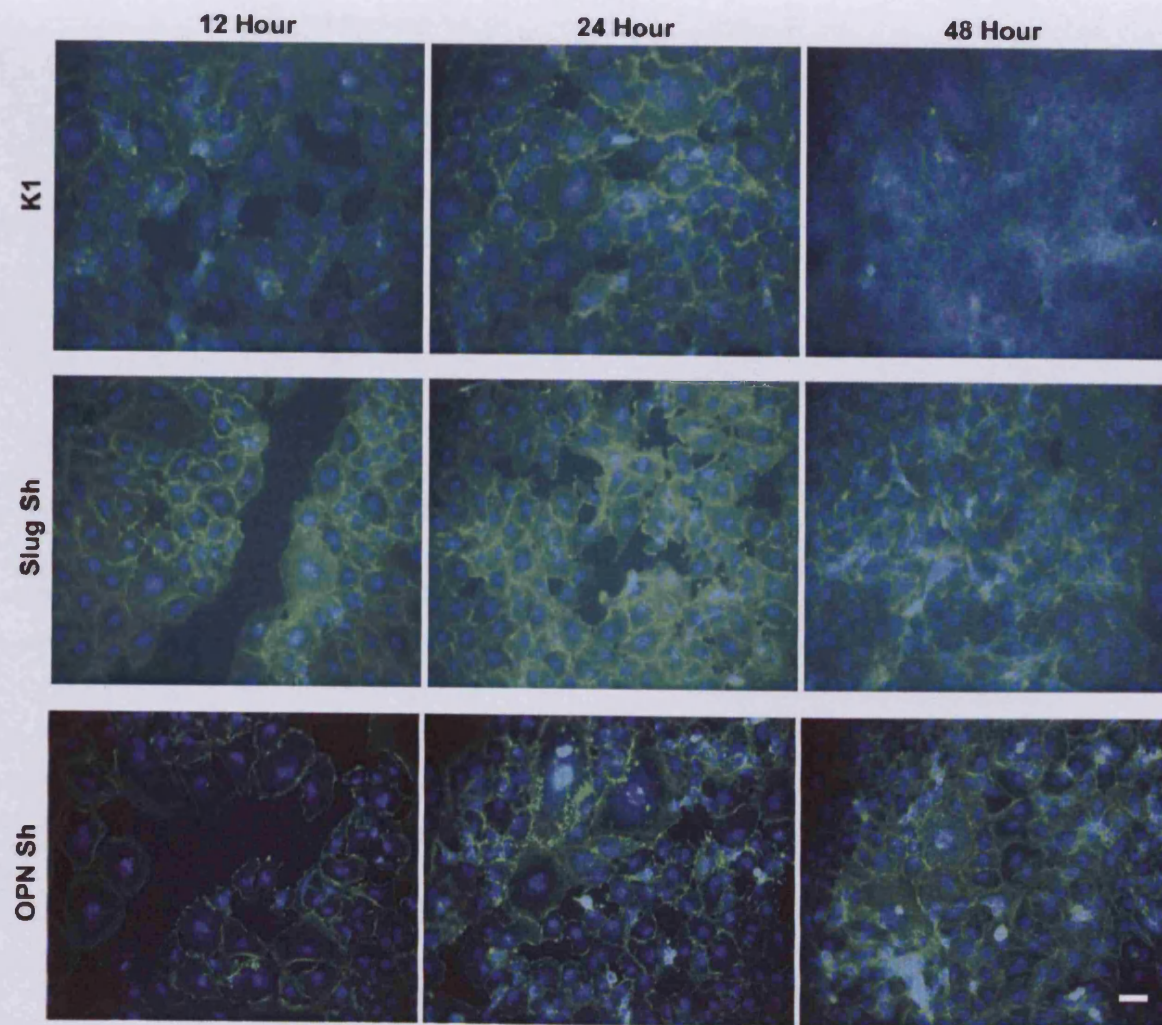
as the wound closes in OPN Sh, the cell body at the leading edge increase in size. This is evident at the 8 hour time point. The cell body of both K1 and Slug Sh at the leading edge did not increase in size like OPN Sh (Figure 35).

This experiment essentially observed cell motility at selected time points. To expand on this, I conducted time-lapse analysis to observe cell motility over a period of time and not just at selected time points, thereby providing further details of possible mechanisms which regulate invasion.



**Figure. 35.** Immunofluorescent timecourse observing wound closure with F-actin staining. Cells were grown to confluence on coverslips, wounded and after 3 hours fixed and stained at selected time points for F-actin (green) with DAPI to observe the nucleus (blue). Red arrows indicate concentration of actin. White arrows indicate the enlargement of cells and lack of actin stress fibres. Scale bar is equal to 50µm, and is applicable to all the images.





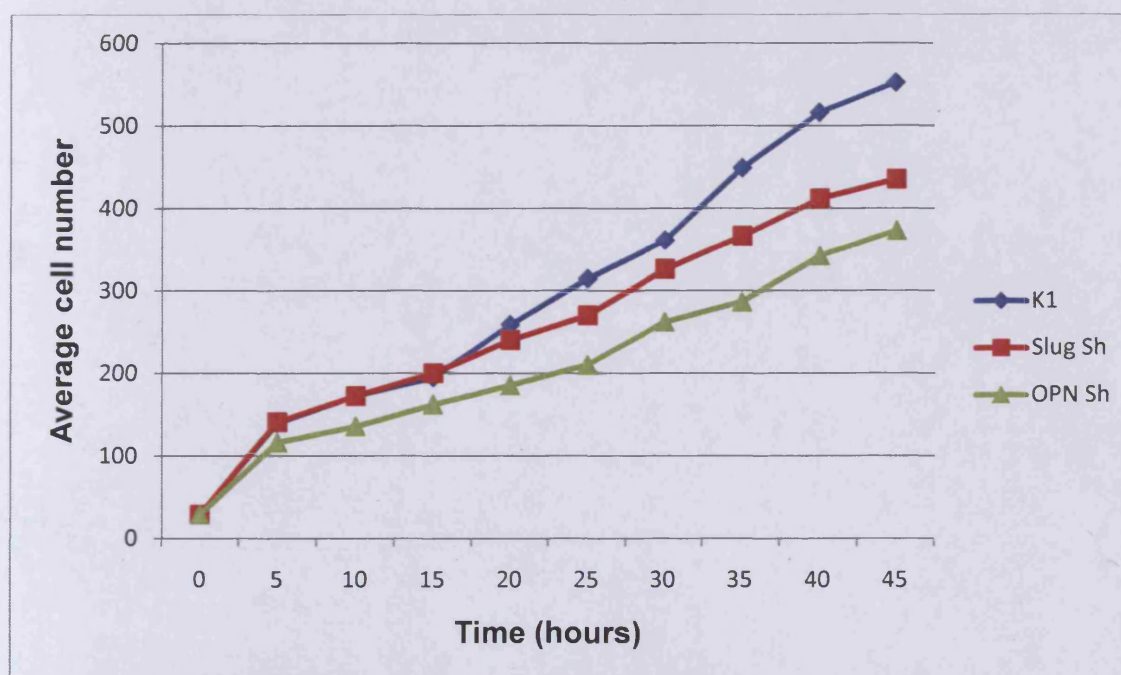
#### **6.6.2.2 Analysis of cell motility**

An integral requirement for transformed tumour cells to invade through the basement membrane and metastasis, not only relies upon adhesion dynamics but also a motile phenotype. The time-lapse capture of tumour cells allows for the analysis of cell movement, motility together with interactions between cells and the cell-substrate contact on a 2-dimensional substrate. To conduct time-lapse analysis, K1, Slug Sh and OPN Sh were used. The time-lapse is useful to study the effect of knocking down expression of Slug and OPN on cell migration, motility and interactions between cells and with the substrate, all of which have never been studied in this manner before.

The time-lapse study was conducted by capturing images of three fields from each cell line every 15 minutes for a period of 48 hours. The cells were plated at a low confluency to account for cell density effects which can attenuate cell migration/motility.

#### 6.6.2.2.1 The effect of Slug and OPN down-regulation on cell growth

The time-lapse camera was used to collect images every 5 hours. The cell number was determined for each image and the data graphically shown. The increase in cell number over time provided an indication of cell proliferation for K1, Slug Sh and OPN Sh. From the analysis, K1 cells displayed the greatest rate of cell number over time totalling ~550 cells over 48 hours, a result which was expectant of a tumour cell line with high invasive capability. The Slug Sh cells had a decreased growth rate at ~410 cells over 48 hours. OPN Sh was further decreased with an even slower growth rate at ~370 cells (Figure 36). These results recapitulate the *in vitro* observations.



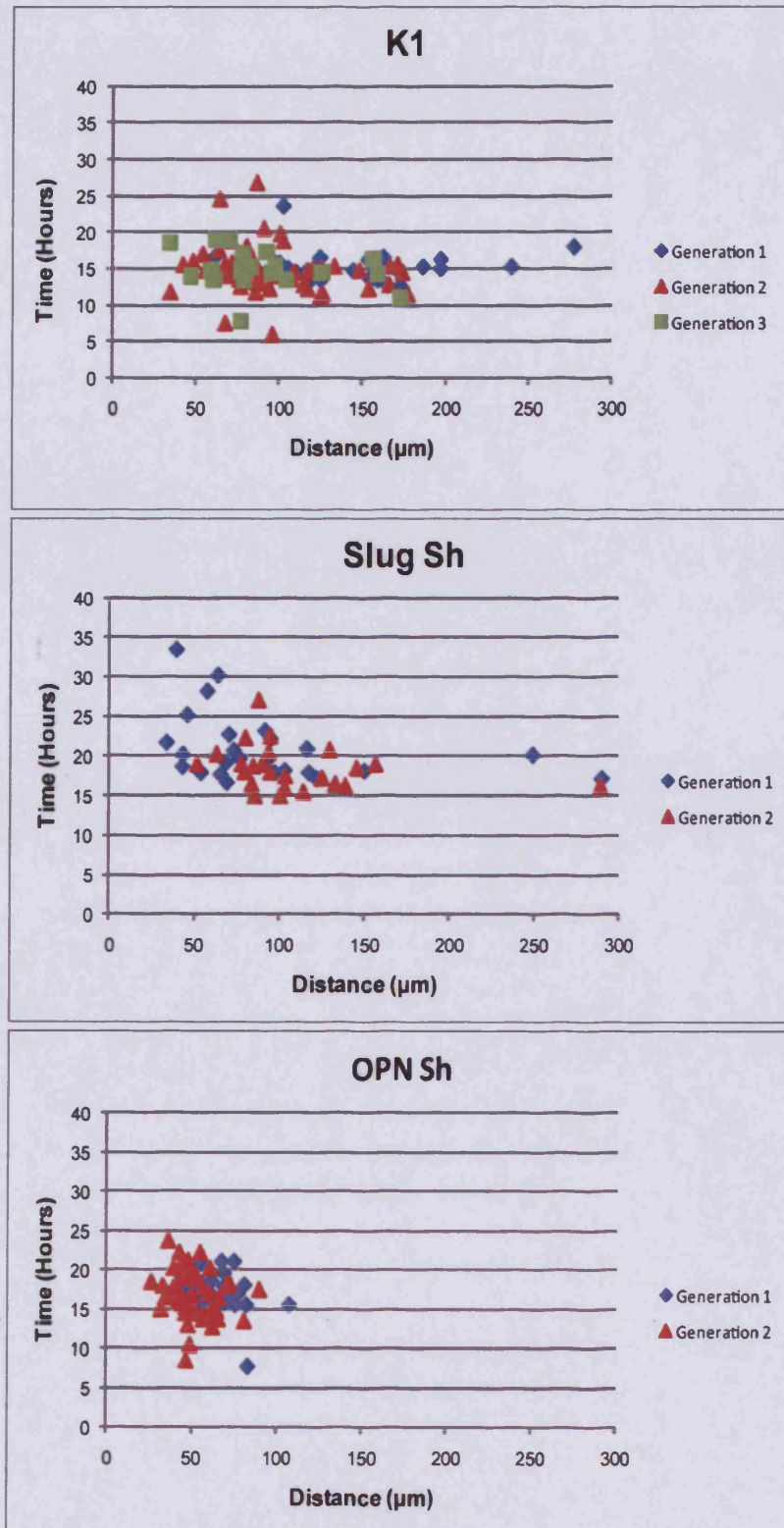
**Figure 36.** Growth curve to measure the number of K1, Slug Sh and OPN Sh cells every 5 hours. The total number of cells on each field was counted every 5 hours and the data averaged to obtain an average cell number per cell line (K1;blue, Slug Sh;red, OPN Sh;green).

Details regarding cell cycle can also be determined using time-lapse, where the time taken between cell division is defined as the 'inter-mitotic time' (IMT). The 'inter-mitotic distance' (IMD) is defined as the total distance travelled in the time between cell division. The inter-mitotic times and distances were defined from a total of 15 cells per cell lines which were tracked using the Metamorph software suite (Molecular devices; <http://www.moleculardevices.com>), which generated up to 3 generations. Cells which could not be tracked in 48 hours were subsequently removed from the analysis.

For every inter-mitotic time track, the corresponding average distance travelled by the cell was determined, and the data graphically summarised by plotting the inter-mitotic times against inter-mitotic distances (Figure 37). This effectively provided a distribution of cell division and distance travelled.

From the results, the K1 cells travelled the greatest distance with a heterogeneous distributed pattern across 3 generations. Interestingly, only the K1 cells had complete tracks in the 3<sup>rd</sup> generation indicating a faster inter-mitotic time compared to both Slug Sh and OPN Sh. Although Slug Sh contained outliers, the overall distribution of the tracks is indicative of a longer mitotic time with a shorter distance travelled. OPN Sh showed the greatest homogeneous distribution compared with K1 cells with distance limited to 125µm, with an inter-mitotic time comparable to Slug Sh.





**Figure 37.** Inter-mitotic time Vs. inter-mitotic distance for each complete generation of K1, Slug Sh and OPN Sh cells. Blue diamonds correspond to the 1<sup>st</sup> generation of cells, Red triangles correspond to 2<sup>nd</sup> generation cells and green squares refer to the 3<sup>rd</sup> generation cells. K1, Slug Sh and OPN Sh are represented as individual graphs.

#### 6.6.2.2.2 Random walk analysis

The movement or motility of cells on a 2D surface can be characterized through random walk analysis, allowing the various distribution patterns (both temporal and spatial) to be mathematically recast. This type of analysis is particularly suited to cellular tracking for both (un)perturbed systems and can inform numerically and analytically assign well-known statistical measures to the individual or population (Selmeczi et al. 2005). In this study, we employ only a few statistical measures to quantify population movement of the differing cell-lines, namely, the asphericity a parameter derived from the radius of gyration tensor and the diffusion coefficient (motility) determined from numerical fitting of the averaged mean square traverse of the population as a function of time (Codling et al. 2008). This particular analysis was conducted by Dr. Martyn Brown, who used the MATLAB programming suite (<http://www.mathworks.co.uk>) to conduct this analysis.

Below, is a brief overview describing how these statistical measures are extracted from the time-lapse data files. From the tracking datasets the X and Y coordinates of individual cells were recorded at each sampling point over the time-course of the experiment. These values were averaged to form the following covariance matrix:

$$T = \begin{pmatrix} \langle x^2 \rangle - \langle x \rangle^2 & \langle xy \rangle - \langle x \rangle \langle y \rangle \\ \langle xy \rangle - \langle x \rangle \langle y \rangle & \langle y^2 \rangle - \langle y \rangle^2 \end{pmatrix}$$

here the angular brackets,  $\langle \rangle$ , refer to an ensemble average.

The matrix in (1) is next diagonalised to reveal the magnitude of the square of the two characteristic orthogonal eigenvalues,  $R_1$ . These values are then comparable to the major and minor axis of an ellipse that describing the average movement of the

cellular population. This analysis thus far gives an overall extent of the random walk but does not account for the shape. The Asphericity parameter,  $A_2$ , is combination of the derived eigenvalues of the covariance matrix, given by the expression:

$$A_d = \sum_{i>j=1}^d \frac{\langle R_i^2 - R_j^2 \rangle}{(d-1)\langle R_i^2 + R_j^2 \rangle}$$

Here,  $d$  refers to the spatial dimension considered, two in this planar study.

The value of  $A_2$  is bounded to the unit interval [0 1]. For a walk which is spherical, the value is 0. For walks which are highly asymmetric, i.e. the  $R_1 \gg R_2$ ,  $A_2$  tends to unity. Therefore, the magnitude of  $A_2$  can give a direct quantitative value for the averaged cellular traverse as a function of time, i.e. generation.

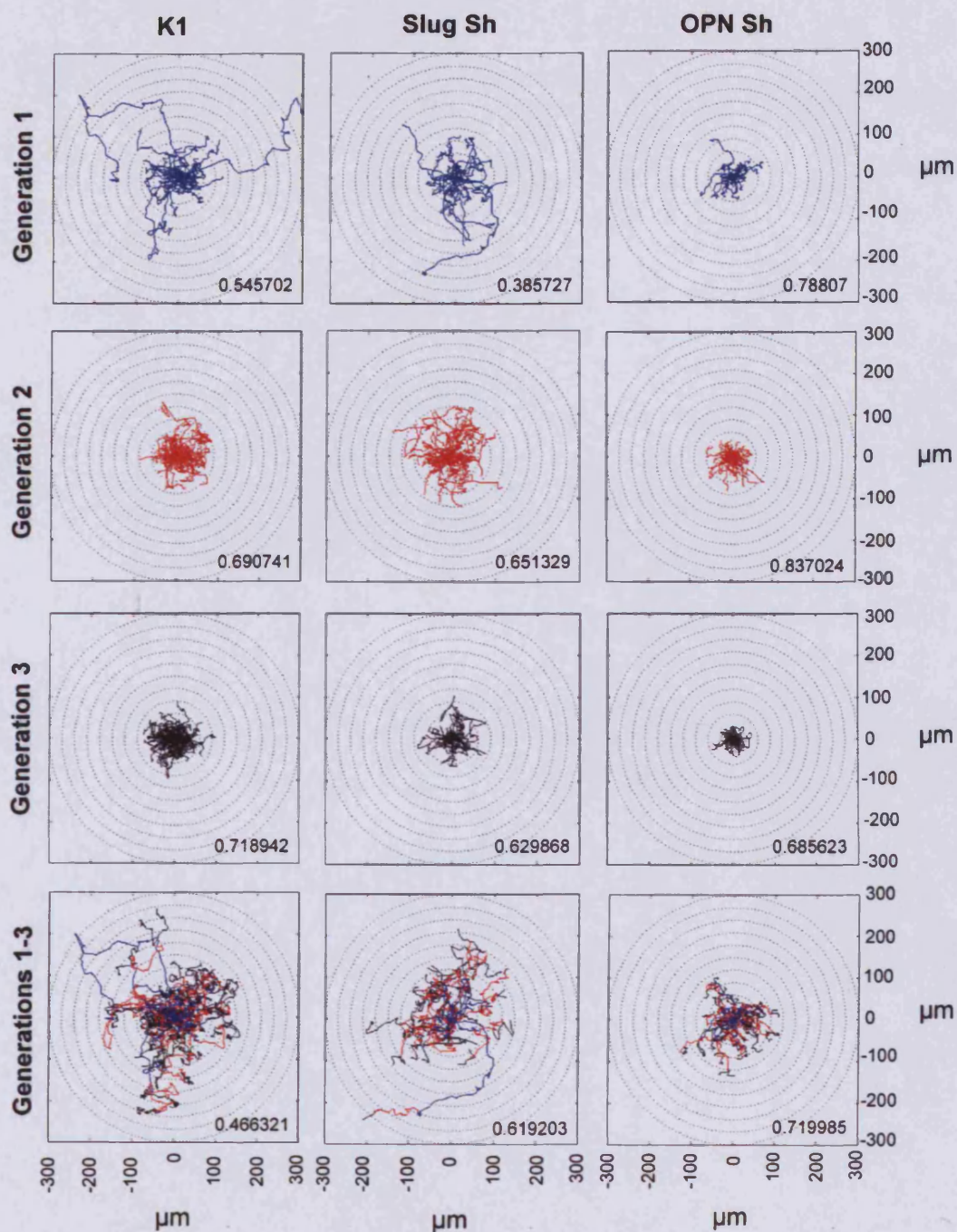
The lesser the value of asphericity is an indication of increased movement from the cells. This is an indicative result due the overlap of the cell tracks at the starting point of 0 $\mu$ m. Individual tracks cannot be sufficiently resolved; therefore the asphericity value cannot be taken as absolute, but merely as an indication of movement. In the 1<sup>st</sup> generation Slug Sh cells surprisingly had the lowest asphericity value which can be visually observed by the extent of the random walk, and this walk is also comparable to the K1 cells. Most interestingly, the OPN Sh cells displayed a highly spherical random walk signified visually by the walk centred on 0 $\mu$ m with an asphericity value of 0.788. The 2<sup>nd</sup> generation is particularly interesting as in general all the cells displayed a spherical random walk compared with the 1<sup>st</sup> generation. The K1 cells had an asphericity value of 0.691, which is highly comparable to Slug Sh cells with a value of 0.651. The OPN Sh cells did not alter their respective random walk pattern with increased asphericity of 0.837. In the 3<sup>rd</sup> generation asphericity was not significantly altered with K1 slightly increasing in asphericity to

0.720 by both Slug Sh and OPN Sh decreasing in asphericity, however visually the random walk seems less random-like (Figure 38).

#### *6.6.2.2.3 Analysis of velocity during cell migration*

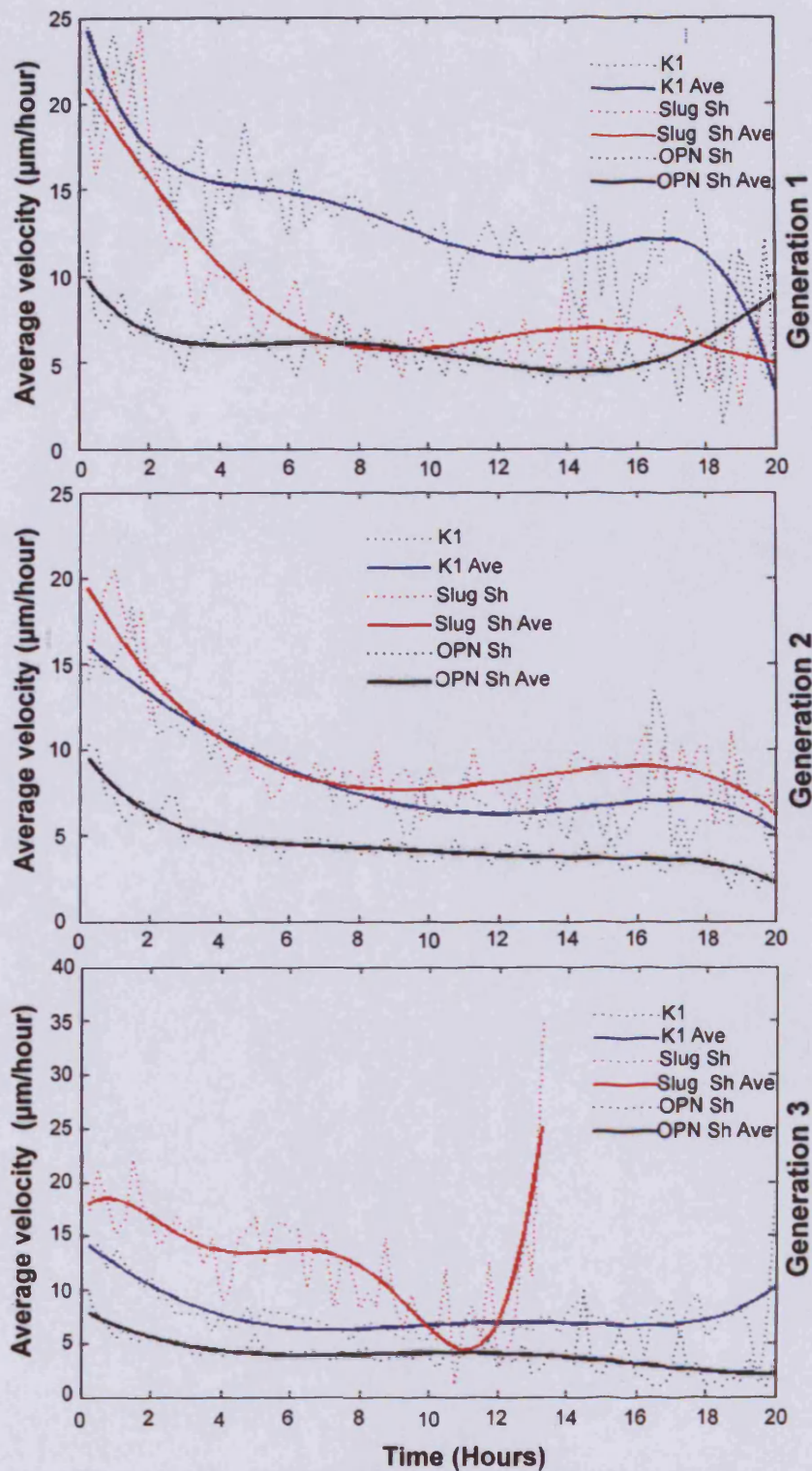
The velocity of the cell can be used as a measure of cell movement, where a slower velocity would equate to less movement, and increased velocity with increased motility. To investigate and analyse cellular velocity, the Metamorph suite collects velocity values during the tracking of cells. The same cells which were tracked for their inter-mitotic times, distance and their random walk were analysed for their average velocity across each generation. Analysis of each generation was conducted separately to investigate the effect of cell proliferation on cell movement. Incomplete tracks were defined as cells which migrated from the field of view. These incomplete tracks were included in this particular analysis.

The results show K1 to be the quickest. Following the pattern of random walk and inter-mitotic distance, Slug Sh displayed decreased velocity with OPN Sh even more so. After the 1<sup>st</sup> generation all the cells had a decreased velocity which may be attributed to cellular proliferation, and lack of space for the cell to move into. Overall the 2<sup>nd</sup> generation velocity decreased compared with the 1<sup>st</sup> generation. The average velocities of K1 and Slug Sh in the 2<sup>nd</sup> generation closely followed each other at a velocity of 10  $\mu\text{m}/\text{hour}$  at 6 hours, whilst OPN Sh remained the slowest at 10  $\mu\text{m}/\text{hour}$ . Surprisingly in the 3<sup>rd</sup> generation, Slug Sh cells had the most velocity, with an exponential increase in the space of 2 hours at the end point of analysis (Figure 39).



**Figure 38.** Generational Random Walk analysis for K1, Slug Sh and OPN Sh. A measure of the asphericity of the random walk is defined by the value given in each panel. The closer the value to 0 indicates a spherical random walk, indicating less movement of the cells. A value closer to 1 indicates more movement.





**Figure 39.** Preliminary experiment displaying the average velocity for each generation of K1, Slug Sh and OPN Sh cells. Each graph represents each generation of 15 cells. Dashed lines represent actual values, whilst solid lines represent average values. K1 is highlighted by blue, Slug Sh as red and OPN Sh as black.

### 6.6.2.3 Summary

Cell motility, an inherent quality of invasion was investigated by time-lapse microscopy. This method is useful in investigating altogether proliferation; cell-cell interactions, cell-substrate interactions and how the cells move across a 2D substrate. By tracking cells from K1, Slug Sh and OPN Sh vital statistics including position, time and speed were determined. Simply determining the amount of cells at each time point effectively determined the cellular growth. Since invasive capability is impaired in both Slug Sh and OPN Sh, an expectant result would be an overall affect on motility. From this investigation, over time, Slug Sh had a decreased cell number with OPN Sh having a further decrease in cell number compared to the PTC cell line, K1. The attenuated cellular growth may be a phenotypic alteration which ultimately affects invasive capability. Although the cellular context is different, in human gliomas Slug was shown to promote cell growth and invasion (Yang et al. 2010). Here, knockdown of Slug decreased cell growth, and is consistent with Yang et al. The mechanism of how cell growth is affected by Slug requires further investigation. OPN Sh displayed a significant decrease in cell number/time, and this finding correlates with other studies, where OPN overexpression is linked to cell growth and tumour invasion (Philip et al. 2001).

Tracking individual cell lineages of K1, Slug Sh and OPN Sh cells allowed the movement and inter-mitotic times to be investigated. This analysis revealed phenotypical alterations which the Slug Sh or OPN Sh induce. The distance Vs. time analysis revealed a wide distribution in K1 cells. For Slug Sh cells, the distance travelled was decreased and there was a further decrease in movement for OPN Sh cells. K1 cells had the fastest inter-mitotic time with the greatest degree of

movement. Slug displayed a decreased inter-mitotic time with a lower degree of distribution of movement/time. OPN knockdown displayed the greatest alteration with a severely restricted movement. Linking with the invasive capacity of the knockdown clones, there is a reduction of movement. The process of movement is integral to invasion, without it cells would not display an invasive capability. This work provided evidence of phenotypic alterations to cell migration and this correlates with the reduction of invasion. Further investigation of individual generations by random walk analysis provided a visual representation of movement. Over the entire time course when cells proliferate, the free space available for cells to move becomes restricted. This means the results from generation 2 onwards are skewed as free space is a limiting variable which is inherent to the experiment. Observation of the 1<sup>st</sup> generation alone provides an experiment devoid from this limiting variable. The determination of asphericity is in itself not an absolutely accurate method. This method does not have a high degree of resolution where cell tracks overlap each other, which therefore affect the asphericity values. The effect of proliferation is also evidenced by the data pertaining to average velocity. As cells divide the free space to move becomes less, and so less movement correlates with reduced velocity. Interestingly OPN Sh displayed the most reduced velocity even though space and proliferation in the 1<sup>st</sup> generation was not a factor affecting velocity. Analysing the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> generation individually allows cell proliferation and cell density to be analysed across each generation and how this impacts cell movement and motility.



## 6.7 Epithelial Mesenchymal Transition

The process of EMT is prominent in tumour invasion. Tumour cells utilise EMT to further their progression in invasion, metastasis and survival. EMT brings about a change in cellular morphology to allow cells to have a migratory phenotype. Alteration of an epithelial cell to have mesenchymal properties requires morphological, architecture, adhesion and migratory changes. Normal epithelial cells form a tight monolayer and their movement is restricted and provide tissues with mechanical rigidity. The presence of cell-cell junctions and adhesive molecules holds the cells tightly together and inhibits the movement of individual cells. In contrast, mesenchymal cells do not possess the junctions and adhesive properties like epithelial cells, but their adhesive properties are less strong allowing for greater motile capacity. They form structures which are irregular in shape and not uniform in structure or density. The morphology of mesenchymal cells is irregular as they possess leading edge polarity (Chambers et al. 2002; Hugo et al. 2007). The alteration from epithelial to mesenchymal cells defines an essential dynamic process for tumour progression. An EMT transition involves a change in the adhesive qualities of the cell, remodelling of the cytoskeleton, and proteolytic degradation of the extracellular matrix.

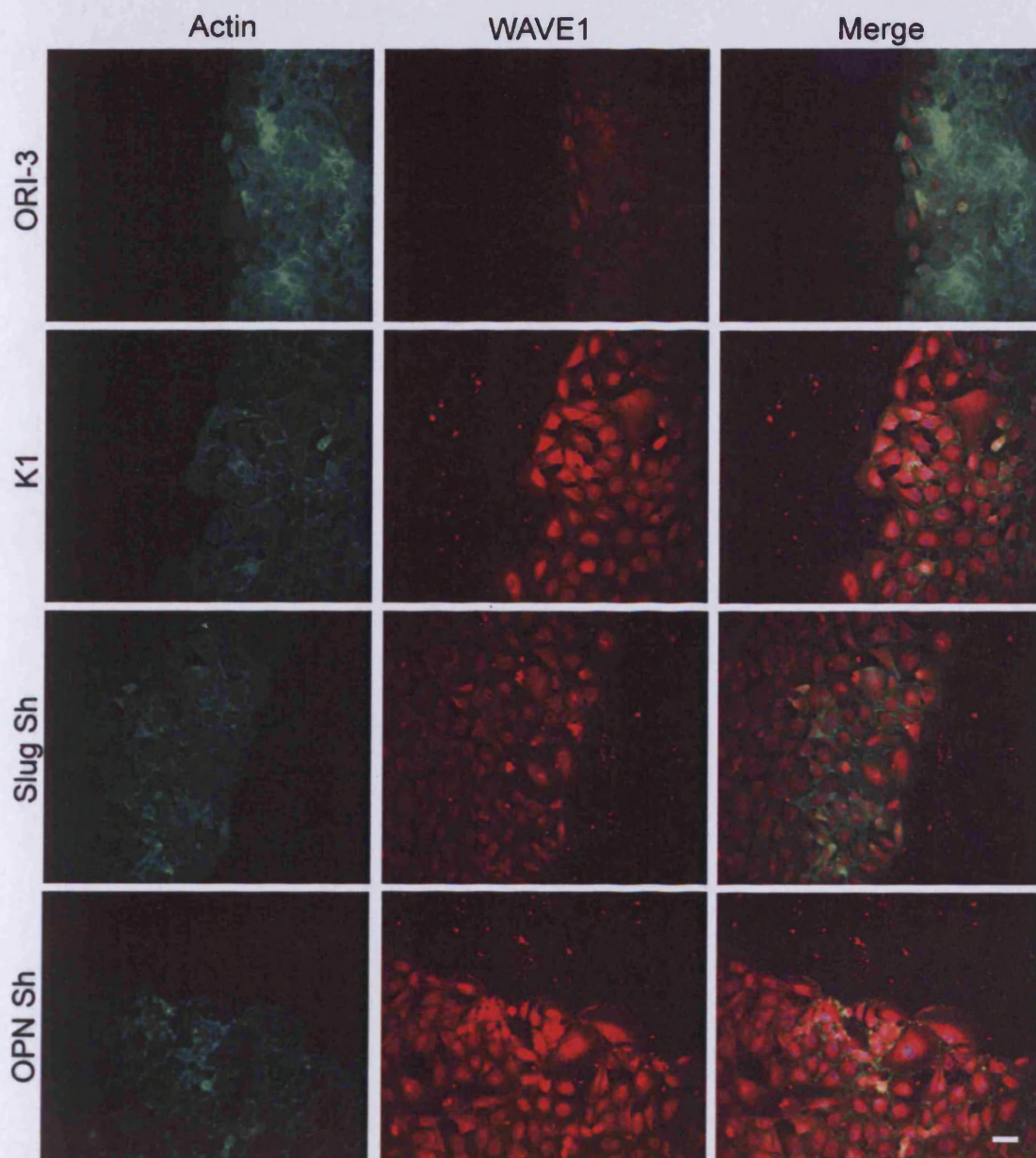
In my investigation I aimed to study if reversal of EMT occurs in Slug Sh and OPN Sh. We hypothesised that down-regulation of Slug and OPN would convey an increased epithelial phenotype in terms of cell motility, cell-cell adhesions and cell-matrix adhesions. This in effect would provide reasoning as to why invasive capability was reduced in Slug Sh and OPN Sh.

### **6.7.1 WAVE and N-WASP in actin-based cell movement**

The actin filament system regulates cell movement. The WASP family of proteins including N-WASP and WAVE1 activate the Arp2/3 complex which mediates the formation of lamellipodia and filopodia. Formation of lamellipodia and filopodia are integral for cell movement. Signalling upstream to WAVE1 and N-WASP are the small GTPases Rho, Rac and cdc42.

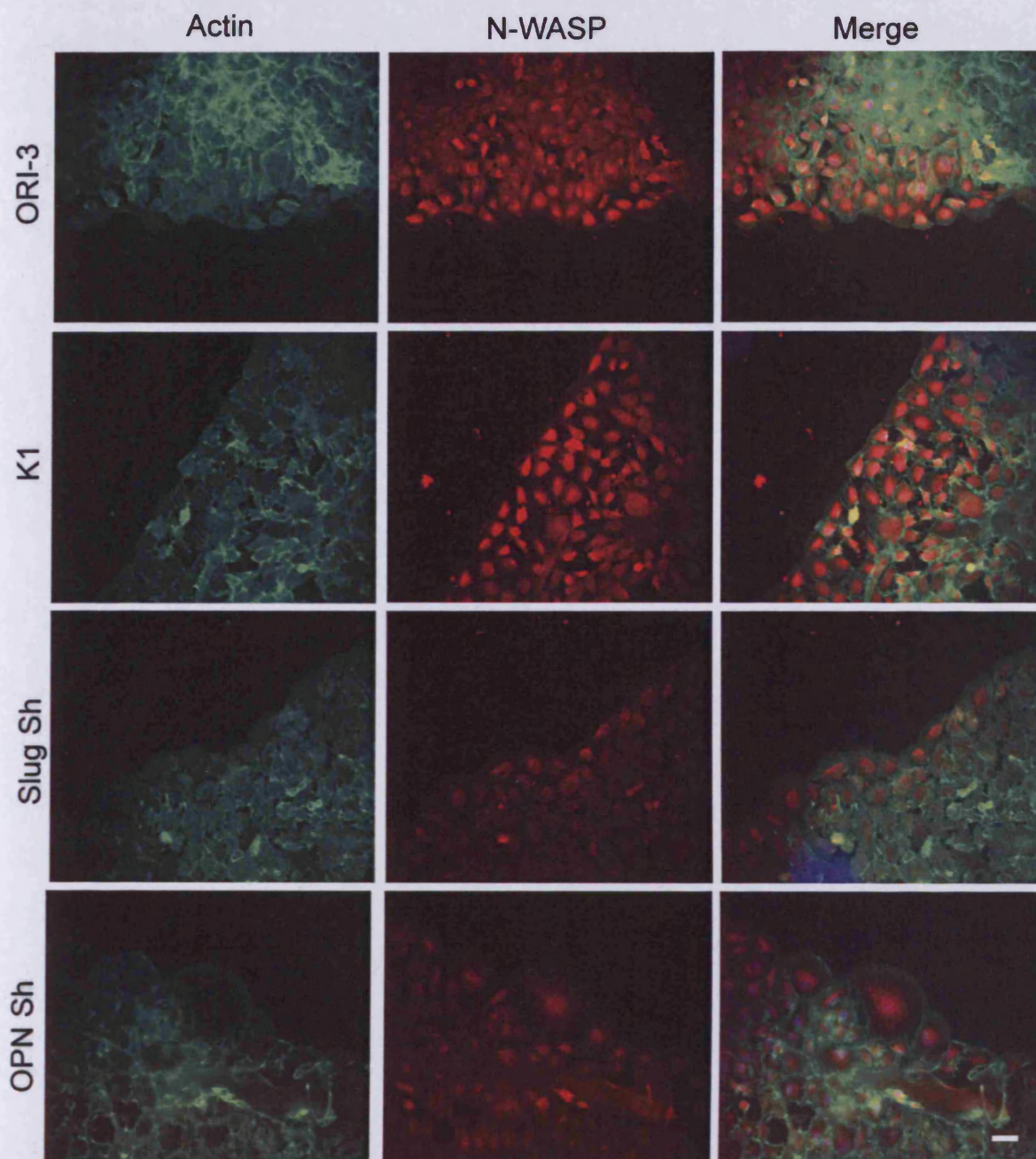
To investigate if this system of actin protrusion formation is affected by the knock-down of Slug and OPN, I immunofluorescently stained cells for WAVE1 and N-WASP. Cells were grown to confluence and wounded. After 3 hours the cells were fixed and probed for WAVE1 and N-WASP using immunofluorescence. This enabled me to visualise the localisation of these proteins.

WAVE1 expression was localised in the nucleus and cytoplasm in ORI-3 cells but at basal levels. In K1 cells, there is a significant increase in signal intensity. Perinuclear staining of WAVE1 was demonstrated in Slug Sh cells. Interestingly OPN Sh cells have an intense nuclear and perinuclear staining of WAVE1. However this may be non-specific artefacts (Figure 40). The N-WASP protein in both the Slug Sh and OPN Sh cells displayed reduced signal intensity in comparison to K1 cells (Figure 41). The localisation of WAVE1 and N-WASP is important as expression at the leading edge of cells would indicate a role in lamellipodia and filopodia mediated cell movement.



**Figure 40.** Immunofluorescent analysis of F-actin and WAVE1 expression in Slug Sh and OPN Sh cells. Cells were stained with F-actin (green) and co-immunostained with WAVE1 (red) followed with counterstaining with DAPI nuclear stain (blue). Fluorescence was observed with a fluorescent microscope. Scale bar is equal to 50 $\mu$ m, and is applicable to all the images.





**Figure 41.** Immunofluorescent analysis of F-actin and N-WASP expression in Slug Sh and OPN Sh cells. Cells were stained with F-actin (green) and co-immunostained with N-WASP (red) followed with counterstaining with DAPI nuclear stain (blue). Fluorescence was observed with a fluorescent microscope. Scale bar is equal to 50 $\mu$ m, and is applicable to all the images.

### 6.7.2 Summary

The actin network and the formation of actin-based protrusions are important for cell migration as they regulate the forward thrust required for cell movement along with the dynamic regulation of cell contacts. The lamellipodia and filopodia are types of membrane protrusions which generate forward thrust facilitating cell movement. Movement involves a process where actin filaments polymerise at the leading edge and depolarise at the rear resulting in a forward motion. Central to the formation of these protrusions are the Rho family of GTPases, Rho, Rac and cdc42. Both cdc42 and Rac regulate the leading edge of the cell by inducing the formation of membrane protrusions including filopodia lamellipodia. Rho regulates the rear of the cell by stimulating the formation of stress fibers, inducing actin-mediated cell contraction (Wu et al. 2004). The formation of new actin filaments produces an actin filament network which pushes the plasma membrane forward.

It appears OPN Sh have an impaired actin network in terms of movement. Since Rho functions to formulate stress fibres, the evidence presented in this section implies Rho activity is impaired in OPN Sh. In OPN Sh, Rac and cdc42 are functional, promoting forward movement at the leading edge of the cell. Impaired Rho activity would perhaps result in reduced formation of actin stress fibres, increased cell attachment to the substrate at the rear of the cell, and reduced contractile movement at the rear of cell. Indeed in this thesis, lack of stress fibres has been observed in OPN Sh. To observe if this is the case conducting a Rho activity assay is a crucial experiment. I attempted this experiment by growing K1, Slug Sh and OPN Sh cells to maximum confluency in tissue culture flasks. I next 'wounded' the cell monolayer by scratching with a pipette tip and this was left to incubate for 30 minutes. I harvested each cell line by scraping and used the Rho binding domain (RBD) of the Rho

effector protein, Rhotekin to determine RhoA activity. The RBD motif has been shown to bind specifically to the GTP-bound form of RhoA which results in a significantly reduced intrinsic and catalytic rate of GTP hydrolysis. Using a GST fusion protein containing the Rhotekin residues 7-89 allowed the "pull-down" the Rhotekin-RBD/Rho-GTP complex with glutathione affinity beads. The amount of activated RhoA is determined by a western blot using a RhoA specific antibody. I could not determine RhoA activity, and my method requires further optimisation for Rho activity to be determined.

Functioning downstream of Rac and cdc42 are WAVE1 and N-WASP. WAVE1 is specifically activated by cdc42. WAVE stabilises the interactions of the lamellipodium to the substrate (Yamazaki et al. 2005). WAVE1-deficient cells result in an increase in extension speed and lamellipodia, with a decrease in the cell-matrix interactions at the lamellipodium.

Immunofluorescent analysis of WAVE1 could only be used as a qualitative measure of activity it implies the upstream activator of WAVE1, cdc42 is impaired. However, further investigations are required to determine if this is the case. If WAVE1 is in fact impaired then the adhesive contacts which formed at the lamellipodium would be reduced, restricting the migratory ability of the cells.

N-WASP is specifically activated by cdc42. Although a qualitative measure, immunofluorescent staining of N-WASP revealed a strong signal in ORI-3 and K1 cells. Interestingly, both the Slug Sh and OPN Sh cells displayed a diffuse signal. This indicates that the activation of the Arp2/3 by N-WASP would be decreased, and the formation of lamellipodia and filopodia impaired.

Functioning downstream to WAVE1 and N-WASP is the Arp2/3 complex (Welch 1999). The Arp2/3 complex localises to the leading edge of cells, and inhibition of

Arp2/3 leads to attenuation of lamellipodia formation (Goley and Welch 2006). Therefore, impairment of the upstream elements WAVE1 and N-WASP would result in decreased Arp2/3 activity, attenuating lamellipodia formation.

Taken together, the N-WASP low signal intensity implies a possible impairment of the actin machinery regulating the leading edge of the cell. In this thesis I sought to address this by observing focal adhesion assembly at the leading edge.

## **6.8 The role of HGF in Slug and OPN-mediated cell migration**

I next investigated the role of HGF in cell invasion using scratch wound analysis with time-lapse capture. HGF stimulation is known to induce OPN expression, and is a contributing factor of tumour growth and metastasis (Ariztia et al. 2003).

To investigate this hypothesis, cells were grown to confluence and a wound inflicted in the monolayer with a p200 pipette tip. The closure of the wound was captured by time-lapse analysis. This directed wound closure experiment was analysed with and without the addition of HGF at a concentration of 1µg/mL in the culture medium. This effectively allowed me to determine if HGF has a role in invasion, and if this role is mediated by Slug or OPN.

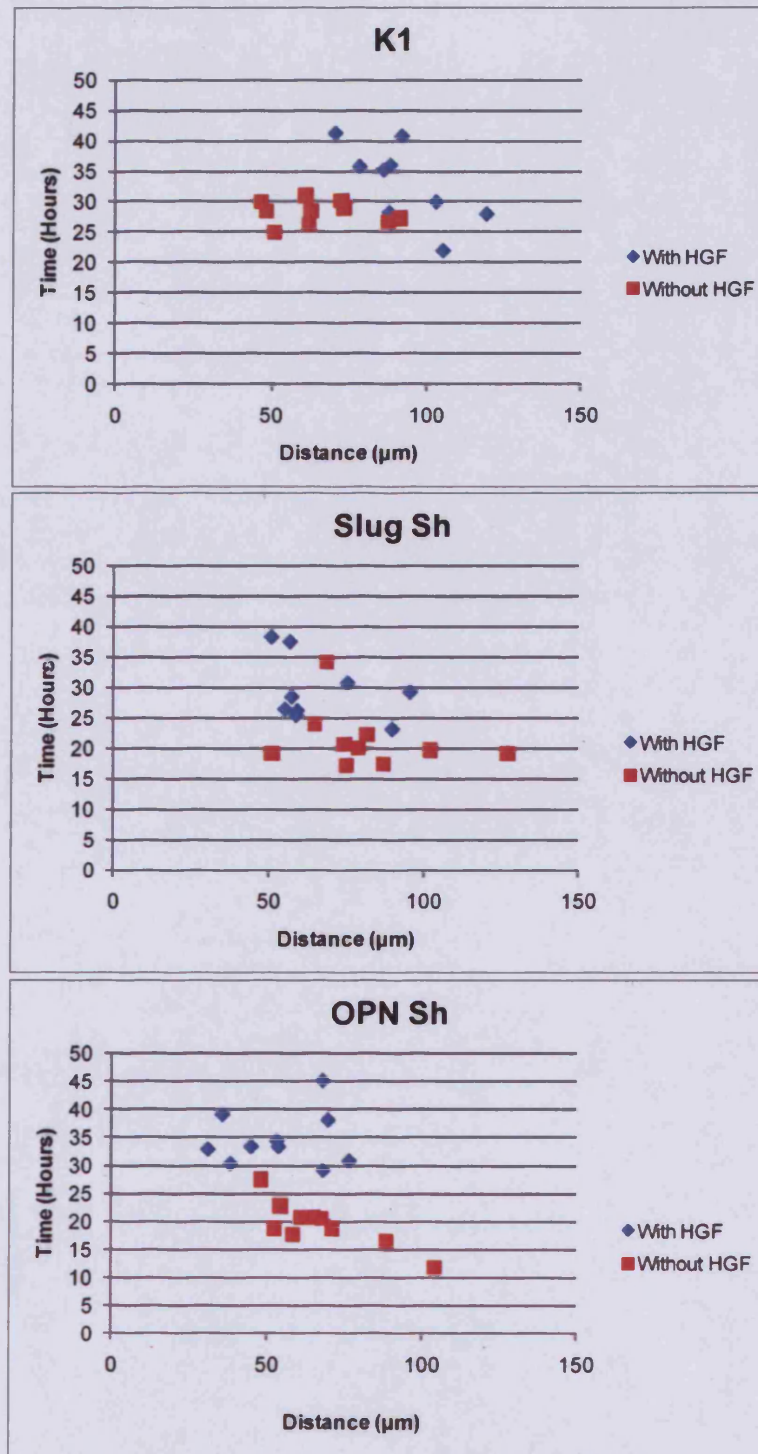
Using the same experimental conditions, the time-lapse capture K1, Slug Sh and OPN Sh cells was conducted and analysed for capacity of the cells to close the inflicted wound over a period of 48 hours. The IMT, IMD and average velocity measurements were collected by tracking cells every 15 minutes using the Metamorph software suite. However, in this instance only the 1<sup>st</sup> generation could be

analysed since all cells failed to complete a 2<sup>nd</sup> generation in the given time frame of 48 hours.

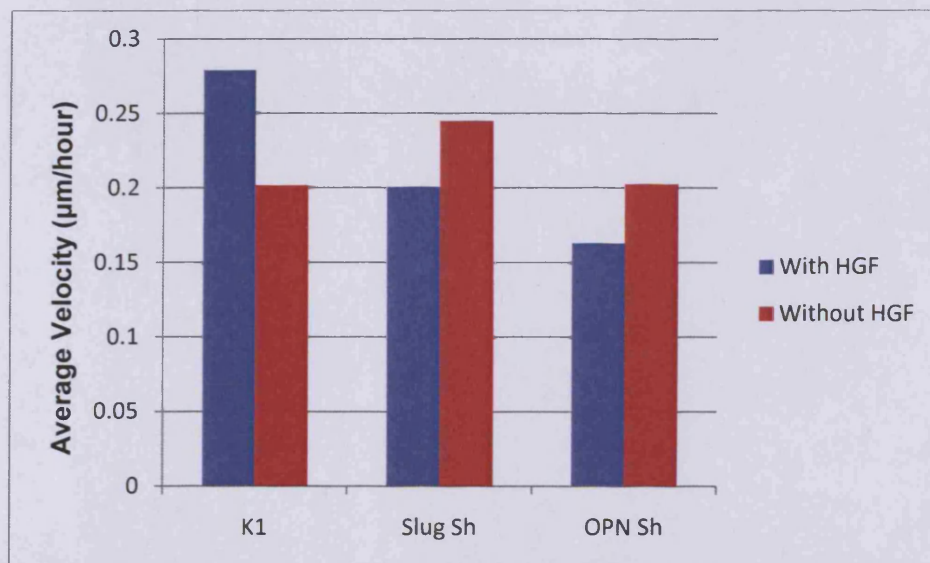
In the K1 cell line the effect of HGF is apparent. Without HGF treatment average IMT for the tracked cells was clustered at 27 hours and at an IMD of 75µm. The addition of HGF resulted in a more heterogeneous pattern with an increase in inter-mitotic time and distance. The Slug Sh displayed a minimal change in IMD with HGF treatment. However, the IMT was increased from 20 hours to 27 hours following HGF treatment. For OPN Sh cells without HGF treatment, the IMT was clustered at ~20 hours. Following treatment with HGF, the IMT increased to ~32 hours, however, the distance travelled did not increase (Figure 42). The average velocity of K1 cells after HGF treatment increased, whilst both Slug Sh and OPN Sh displayed a decrease in velocity following HGF treatment (Figure 43). This velocity data does not correlate with the previous results (Figure 39), where Slug Sh and OPN Sh had reduced velocity compared to K1 cells. This may be due to conducting the HGF-induced migration in a scratch wound assay. In addition, differences within the cells may be more difficult to visualise in this setting.

Taken together, this indicates a probable role of HGF in Slug and OPN mediated invasion, since the knockdown of Slug and OPN did not alter IMD or velocity like K1 after HGF treatment. To further this investigation, I sought to analyse the mRNA content of CD44v, and surface localisation of CD44 as CD44 has a documented role in HGF mediated invasion (Ariztia et al. 2003).





**Figure 42.** Preliminary experiment displaying Inter-mitotic time Vs. inter-mitotic distance for K1, Slug Sh and OPN Sh cells in a wound closure experiment. Cells were treated with and without HGF. Cells were tracked for 1 complete generation, with the IMD and IMT calculated. Blue diamonds correspond to cells with HGF treatment and red squares correspond to cells without HGF treatment. K1, Slug Sh and OPN Sh are represented as individual graphs.



**Figure 43.** Average velocity for the 1<sup>st</sup> generation of K1, Slug Sh and OPN Sh cells in a wound closure experiment. Cells were treated with HGF (1µg/mL) and without HGF. Blue bars correspond to cells with HGF treatment and red bars correspond to cells without HGF treatment.

### **6.8.1 Assessment of CD44v and the implication for HGF-mediated invasion**

Overexpression of OPN induces an increased migratory phenotype in epithelial cells, with the cMet-hepatocyte growth factor (HGF) system playing an integral role. The Met tyrosine kinase receptor is predominantly expressed on epithelial and endothelial cells, its ligand HGF is produced by motile mesenchymal cells. The interaction leads to a metastatic-invasive phenotype promoting EMT (Ariztia et al. 2003). HGF, *in vitro* is a potent stimulus for cell growth, motility and morphogenesis (Birchmeier and Gherardi 1998). OPN is a transcriptional target of HGF and CD44v can bind OPN with cooperation from  $\beta 1$  integrins to mediate HGF-mediated tumour invasion and metastasis (Orian-Rousseau et al. 2002). From the directed cell motility analysis there is an implication role for Slug and OPN in HGF-mediated cell migration.

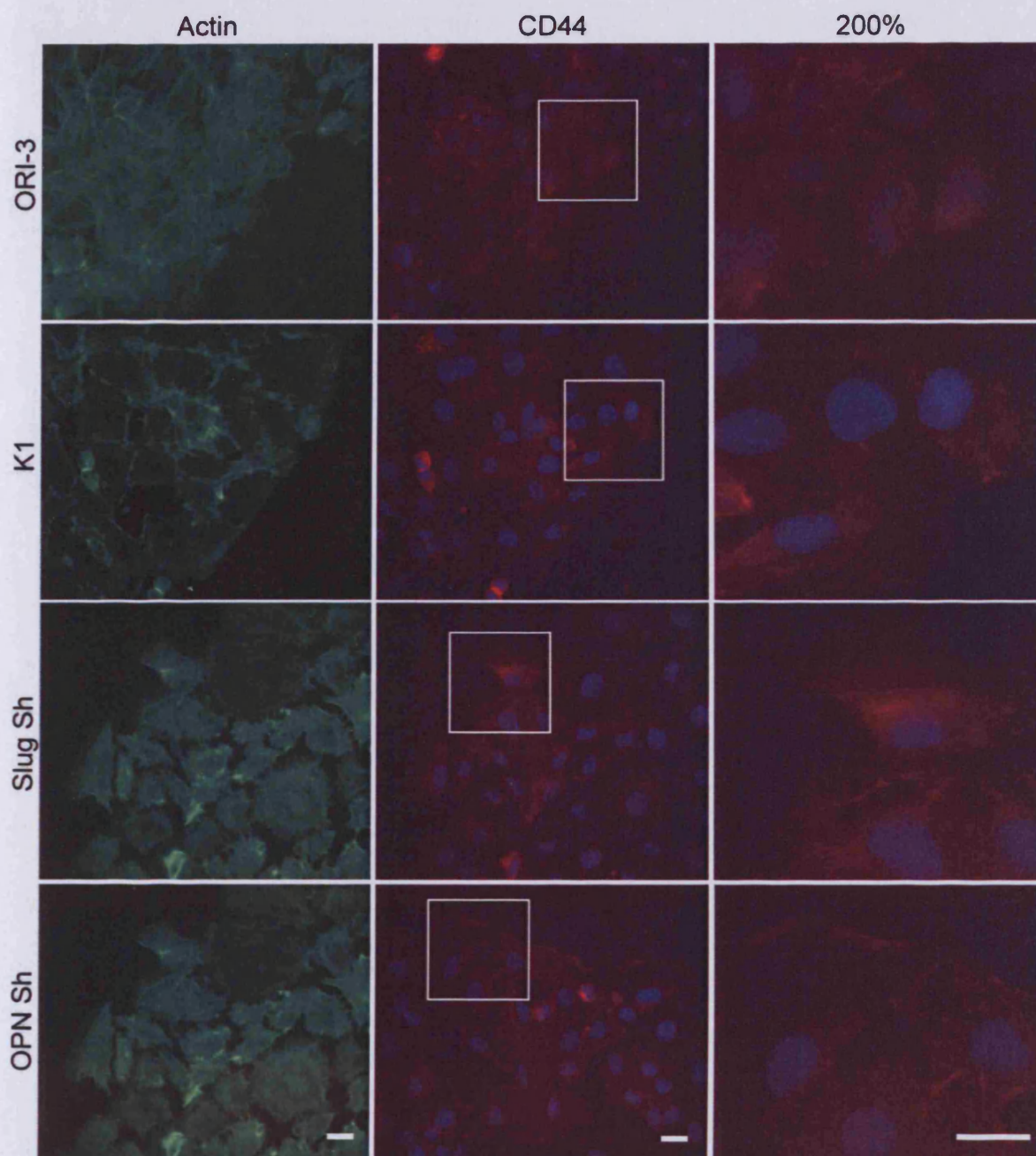
To investigate the role of CD44v in HGF mediated invasion, the mRNA content of CD44 isoforms was determined by RT-PCR. The isoforms analysed included CD44s, CD44v6, CD44v7 and CD44v10. To do this I extracted total RNA from sub-confluent dishes and using the process of RT-PCR. Specific fragments of CD44 standard (CD44S), CD44v6, v7 and v10 were amplified with specific primers. MCF7 acted as a negative control for the experiment, whilst ORI-3 acted as a normal thyroid epithelium negative control. CD44s expression was highlighted in all cell lines, but only the negative controls, MCF7 and ORI-3 displayed decreased mRNA content. The other cell lines displayed equal amounts of the CD44S mRNA. There is clear increased expression of CD44v6, v7 and v10 mRNA in K1 cells. This indicates a potential role of CD44 variants in invasive ability and tumour progression in this cell line and therefore papillary carcinoma. In Slug Sh and OPN Sh a reduction of CD44 variant mRNA is observed (Figure 44).

Observing CD44 surface expression with immunofluorescent analysis, the results show continuous localisation at the leading edge in OPN Sh. Cells at the leading edge of Slug Sh displayed no specific CD44 localisation. In ORI-3 cells, the surface expression of CD44 was punctate, with no localisation at the leading edge. K1 cells displayed a similar CD44 localisation to ORI-3 cells, however, with a more diffuse surface expression (Figure 45). This indicates a possible OPN-CD44 role in regulating motile/invasive behaviour.



**Figure 44.** Analysis of CD44 variants in Slug Sh and OPN Sh cells. RT-PCR was conducted using specific primers directed towards individual isoforms of CD44S, CD44v6, CD44v7 and CD44v10 in 3 independent experiments. GAPDH was utilised as a loading control. MCF7 and ORI-3 cells acted as negative controls and MDAMB231 and K1 cells as positive controls.





**Figure 45.** Surface expression of CD44 (all isoforms) in Slug Sh and OPN Sh cells. Cells were grown to confluence on coverslips wounded for 3 hours and fixed with paraformaldehyde. Cells were stained with F-actin (green), CD44 (red) and counterstained with DAPI nuclear stain (blue). Cells were non-permeabilised to visualise surface expression of CD44. Fluorescence was observed with a fluorescent microscope. Images are presented as F-actin, CD44 merged with DAPI for each cell line analysed. Areas were selected (signified by white squares) and enlarged by 200% to provide further intricate detail. Scale bars are equal to 50µm, and each scale bar is applicable to its column.

### 6.8.2 Summary

One of the mechanisms in which OPN confers a migratory phenotype in tumour cells is through the cMet-HGF system. OPN is prominent in HGF-dependent tumourigenesis, and inhibition of OPN expression attenuates the HGF-invasiveness demonstrated *in vitro* and *in vivo* (Medico et al. 2001; Ariztia et al. 2003). It appears that both OPN and Slug have roles in HGF-mediated invasion. With HGF treatment there was an increase in inter-mitotic time, however, there was no increase in distance travelled. This means the expression of Slug and OPN is important for HGF mediated invasion.

To investigate this further I assessed the status of CD44v. CD44v are able to amplify HGF induced invasion (Smith et al. 1999). OPN can interact with CD44v in conjunction with  $\beta 1$  integrins to induce receptor pair clustering (van der Voort et al. 1999). The CD44v6, CD44v7 and CD44v10 have been shown to bind to OPN to enhance the migratory phenotype of rat fibrosarcoma cells (Katagiri et al. 1999).

Observing CD44s mRNA revealed no alteration in Slug Sh and OPN Sh. This means this isoform is not regulated by either Slug or OPN and has no role in PTC tumour invasion. However, the mRNA content of CD44v6, CD44v7 and CD44v10 were possibly decreased in both Slug Sh and OPN Sh. This implies there is direct regulation by Slug and OPN. Based on literature, this evidence links with the HGF treatment since HGF-mediated invasion may occur via CD44v signalling. Down-regulation of Slug and OPN directly regulates CD44v which impairs HGF-mediated invasion.

A functional interaction between CD44v and Slug was established in a manner that is dependent on CD44v and partially dependent on NF $\kappa$ B and Slug to induce cell

invasion, stimulates by high molecular weight hyaloronic acid, the natural ligand for CD44 (Craig et al. 2009). This is the only study to correlate CD44 and Slug with invasion. Consistent with the evidence presented here, this interaction depends on variant isoforms of CD44, and not CD44s.

Since HGF-mediated invasion by OPN requires CD44v and  $\beta$ 1 integrins, the down-regulation of CD44v mRNA would mean that  $\beta$ 1 integrins are not sufficient to induce HGF-mediated invasion.

Observation of endogenous CD44 by immunofluorescence indicated role of OPN in regulating CD44 surface expression. In melanoma, OPN expression increases CD44 surface expression together with MMP-2 secretion (Samanna et al. 2006). A change in surface expression of CD44 in OPN Sh implies a role of OPN mediating CD44 localisation. This may relate to an invasive mechanism regulated by OPN. It would be interesting to observe surface expression of specific CD44v on the cell surface and possible co-localisation with  $\beta$ 1 integrins. These investigations will elucidate the interacting role of OPN and CD44v in HGF stimulated invasion.

Since OPN Sh has an impaired migratory ability, I sought to investigate the formation of focal adhesions during migration. Focal adhesions are essential to the movement of the cell, the dynamic regulation of adhesions is integral to cell migration and invasive behaviour.

## **6.9 Analysis of focal adhesion formation during cell migration**

In addition to cell-cell interactions, cell-matrix interactions also have a prominent role in EMT. The cell matrix interactions are important for mediating the protractive force required for cellular movement (Harburger and Calderwood 2009). This is conducted through cell-matrix interactions which form, disassemble or mature at the leading edge of the cell and disassemble at the rear of the cell. The adaptor protein vinculin is expressed at focal adhesions (Harburger and Calderwood 2009). The amount and size of focal adhesions is an indication of the motile state of cells, where smaller a fewer adhesions are present on more motile cells.

### ***6.9.1 Analysis of Vinculin and Cortactin***

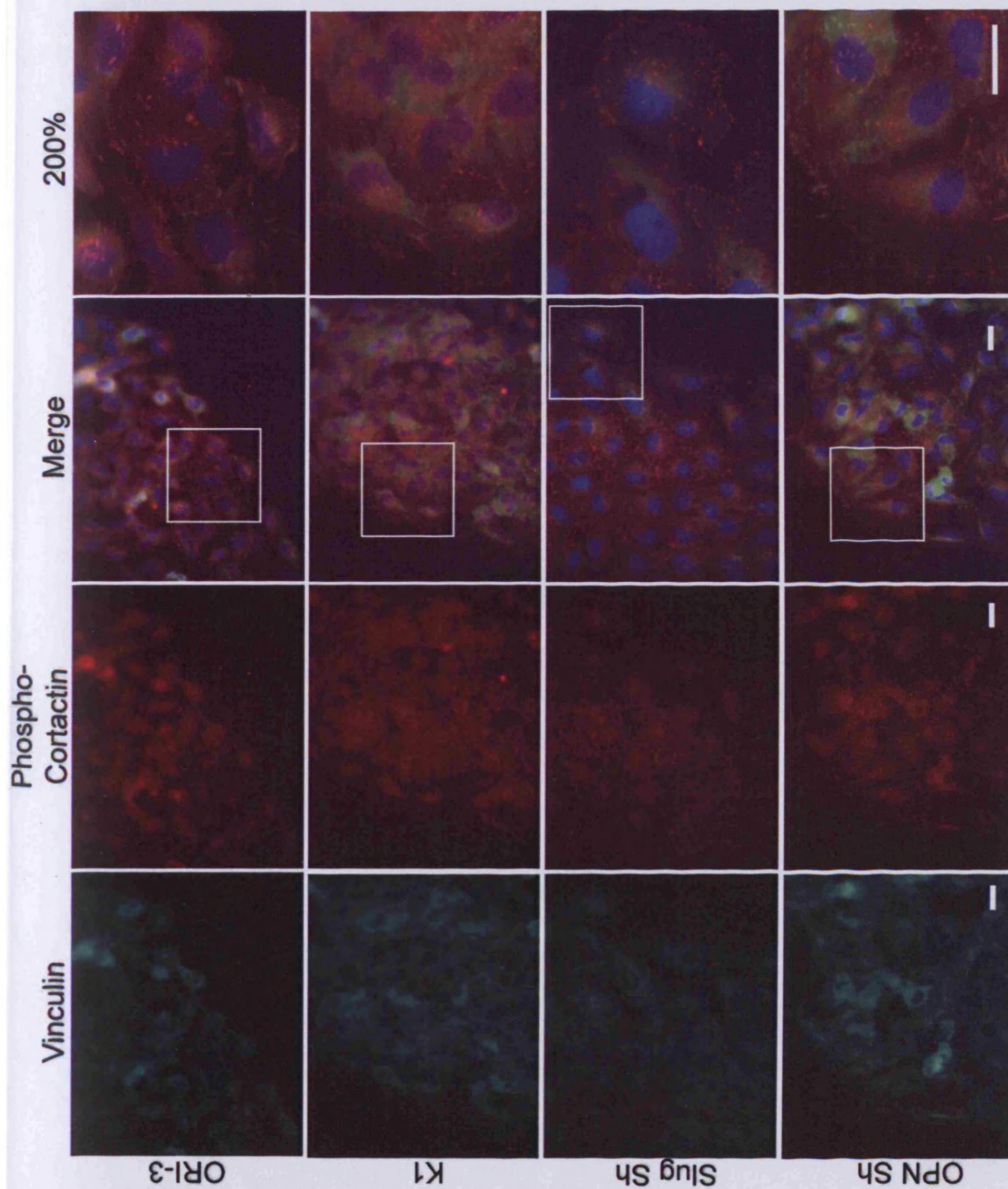
I aimed to assess the localisation of focal adhesions by conferring a migratory phenotype. The localisation and size of the focal contacts are important when considering motile behaviour. The analysis of focal contacts in Slug Sh and OPN Sh may elucidate the mechanism which regulates the reduction in invasive behaviour.

Cells were grown to confluency on coverslips and wounded. After 3 hours incubation, cells were fixed and immunofluorescently stained with specific antibodies. To observe focal adhesions, cells were co-immunostained for vinculin, an adaptor protein present at focal adhesions, along with phosphorylated-cortactin observe the focal adhesions at the leading edge. Cortactin functions at the focal adhesive sites at the leading edge following the formation of lamellipodia. Immunofluorescent staining of vinculin in ORI-3 cells showed distinct focal adhesive sites distinguished by the intense green signal. These sites are pronounced and



substantial with a particular localisation at the leading edge of cells, whilst in K1 cells the vinculin staining is punctate, which may indicate smaller focal adhesions. In addition, the staining of vinculin is prominent at the leading edge of the wound. Slug Sh cells have a slight increase in the size of focal adhesion as indicated by the staining of vinculin, but remain localised around the leading edge of the cells. OPN Sh cells follow a different pattern with large focal adhesions reminiscent of ORI-3 cells. Interestingly, these are localised around the cell membranes of a majority of cells. The expression of cortactin by definition should localise at focal adhesions at the leading edge present in migratory cells. Cortactin expression was localised to the outer edge of cells for the K1 cell line and Slug Sh. There was decreased cortactin expression for the OPN Sh and ORI-3 cell line (Figure 46).

**Figure 46.** Co-immunostaining to detect vinculin and phosphorylated-cortactin expression in ORI-3, K1, Slug Sh and OPN Sh cells. Cells were grown to confluence on coverslips wounded for 3 hours and fixed with paraformaldehyde. Cells were co-stained with vinculin (green) and phosphorylated-cortactin followed by counterstaining with the DAPI nuclear stain (blue). Fluorescence was observed with a fluorescent microscope with an x40 objective lense. Images are presented as vinculin alone (green), phosphorylated-cortactin (red), merged vinculin and phosphorylated-cortactin with DAPI, and selected areas enlarged by 200% (signified by white boxes). Scale Bars indicate 50 $\mu$ m and are applicable to all images.



### **6.9.2 The regulation of Integrin $\beta$ 3 localisation by Slug and OPN**

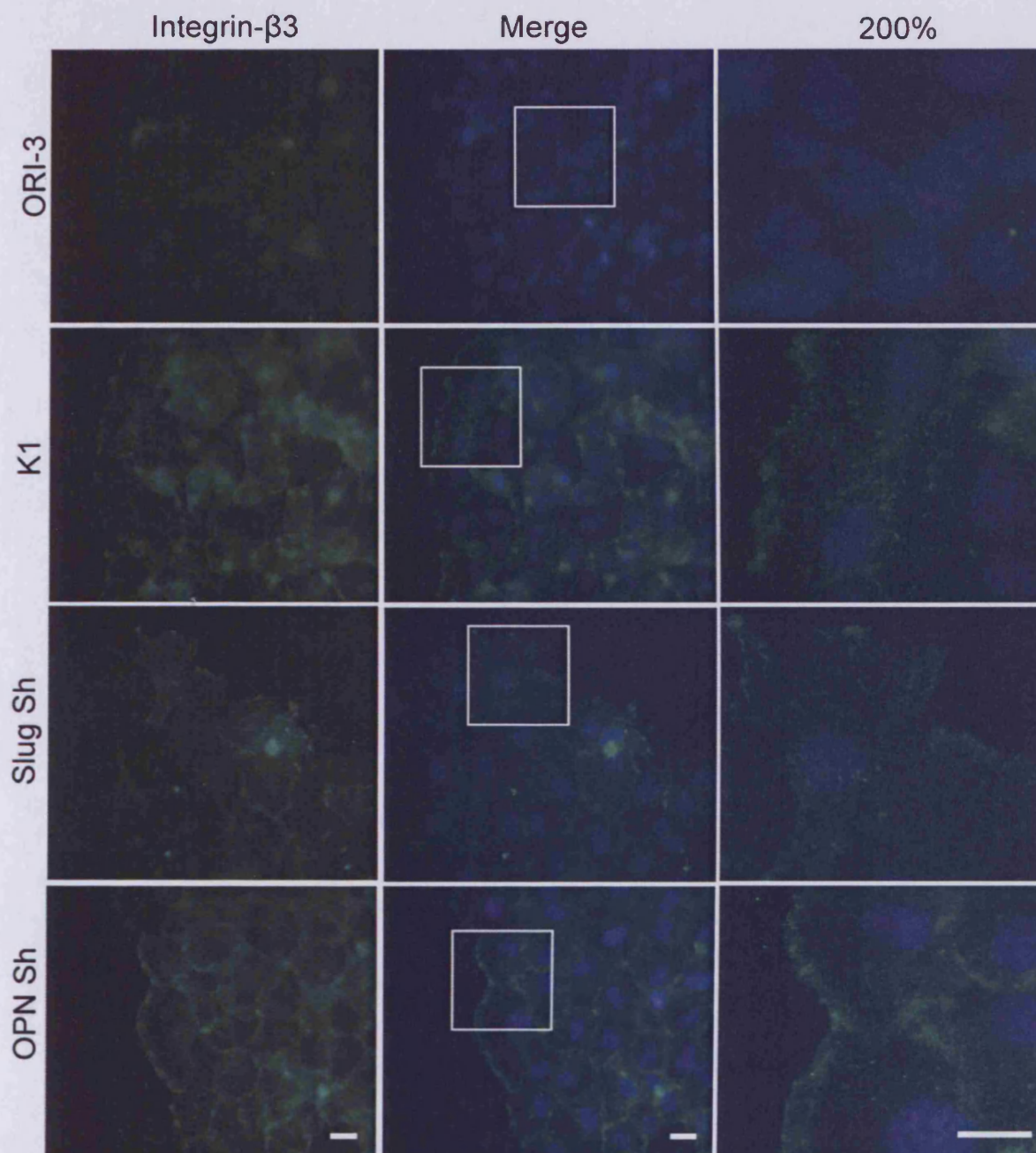
Integrin $\beta$ 3 is important for cell proliferation, survival, invasion and migration (Kiosses et al. 2001). In migrating fibroblasts, Integrin  $\alpha\beta$ 3 is localised at the leading edge of migrating cells over FN and vitronectin ECM components (Woods et al. 2004). This localisation is due to Rac activation which recruits integrin $\beta$ 3 (Danen et al. 2005). Overexpression of integrin  $\alpha\beta$ 3 increases Rac activity and lamellipodium formation, increasing actin stress fibre formation.

To investigate this I immunofluorescently probed integrin $\beta$ 3 to observe if Slug or OPN have roles in mediating integrin $\beta$ 3 localisation. This was conducted on coverslips where cells were grown to confluency, wounded and after 3 hours fixed and immunofluorescent stained.

The immunofluorescence analysis of integrin $\beta$ 3 in ORI-3 cells failed to detect integrin $\beta$ 3. At the leading edge of K1 cells, integrin $\beta$ 3 was localised at the cell periphery and within the cytoplasm. The localisation was focused at points along the leading edge. This indicates probable formation of lamellipodia at these points (Figure 47).

In Slug Sh localisation of integrin $\beta$ 3 at the leading edge was focused at points, however, more sparsely than K1 cells. OPN Sh displayed increased localisation of integrin $\beta$ 3 at the leading edge with no concentrated points of staining, which was observed in Slug Sh and K1 cells (Figure 47). This implies a lack of movement, with no formation of actin protrusions. The continuous staining may relate to the lack of stress fibres which would impair the acto-myosin contractility required for the retraction of the rear.





**Figure 47.** Integrin $\beta$ 3 localisation in ORI-3, K1, Slug Sh and OPN Sh cells. Cells were grown to confluence on coverslips wounded for 3 hours and fixed with paraformaldehyde. Cells were stained with integrin  $\beta$ 3 (green) and counterstained with DAPI nuclear stain (blue). Fluorescence was observed with a fluorescent microscope. Images are presented as integrin  $\beta$ 3 alone, integrin  $\beta$ 3 merged with DAPI and an area enlarged by 200%. Scale bars are equal to 50 $\mu$ m, and each scale bar is applicable to its column.

### 6.9.3 Summary

Directional movement of cells requires the formation of membrane protrusions coupled with cell adhesion. Integrins are central and function as a bridge between the ECM and the actin cytoskeleton of the cell. Integrins with additional proteins form a multi-protein complex to form cell-matrix adhesions (Wiesner et al. 2005). These adhesions are formulated around activated integrin clusters, and these are essential for cell motility and migration. The adhesions are able to mature and disassemble allowing for dynamic regulation of the cell-matrix interaction.

For migration to occur, adhesion structures must form, disassemble or mature at the leading edge of the cell and disassemble at the rear of the cell for the cell to conduct cell migration. Early adhesions mature into focal complexes localised at the border between lamellipodia and lamella. These mature further into focal adhesions with the recruitment of vinculin, a adaptor protein which is always contained within the focal adhesion of the cell (Galbraith et al. 2002). These dynamic structures are found at the cell periphery and in more central areas and are associated with actin stress fibres (Wozniak et al. 2004).

Cells with depleted amounts of vinculin have decreased adhesion and increased migratory ability with smaller and fewer adhesive contacts (Saunders et al. 2006). In addition, vinculin-null cells exhibit increased motility with fewer and smaller focal adhesions, a phenotype prevalent in motile cells (Subauste et al. 2004).

This recapitulates the situation in my investigation. ORI-3 cells display decreased migratory ability, and possess larger focal contacts. K1 cells are highly invasive and possess smaller contacts localised at the leading edge. Slug Sh, revealed localisation of vinculin around the cell membrane and OPN Sh displayed fewer

contacts containing vinculin, and these were situated around the cell membrane. Since vinculin can be present in early focal contacts, further analysis is required to elucidate the type of adhesion present.

Cortactin is an actin filament binding protein which localises at the leading edge of lamellipodia (Uruno et al. 2001). It mediates the activation of Arp2/3 in a synergistic manner with N-WASP. It is therefore implicated in the formation of lamellipodia, and is associated with a migratory phenotype. A reduction in adhesions containing cortactin implies the activity of Arp2/3 would be impaired in OPN Sh. It also indicates the formation of lamellipodia is restricted. This reduced lamellipodia formation may explain the reduced migratory phenotype and subsequent invasive capability of OPN Sh. In Slug Sh, there was not any significant alteration in integrin $\beta$ 3 localisation compared to K1. This indicates that in Slug Sh the lamellipodia formation is functioning in migration.

The localisation of integrin $\beta$ 3 was assessed to see if either Slug or OPN down-regulation impacts the function of integrin $\beta$ 3. In migrating cells, integrin $\beta$ 3 is localised to the cell edges (Kiosses et al. 2001).

The immunofluorescence analysis showed integrin $\beta$ 3 localisation concentrated at points along the leading edge in K1 cells. In Slug Sh, localisation at the leading edge was also concentrated at points of the leading edge, but more sparsely. In addition, in OPN Sh there was intense localisation along the entire leading edge. This evidence indicates a potential role of integrin $\beta$ 3 in regulating migration and invasion.

This is particularly interesting since integrin $\beta$ 3 is a ligand for OPN. Integrin $\beta$ 3 may act in a cooperative manner with other integrins or CD44 to exert OPN dependent migration. However, since the down-regulation of OPN did not impact the localisation of integrin $\beta$ 3 it appears OPN-dependent cell migration does not involve integrin $\beta$ 3.

Integrin $\beta$ 3 also has a prominent role in focal adhesion assembly, and the intense staining concentrated at the leading edges may in fact be focal adhesions. In K1 cells which are highly motile, there needs to be a dynamic regulation of adhesive structures at the leading edge. Knockdown of slug may have reduced the number of focal adhesions, attenuating cell migration. The continuous localisation of integrin $\beta$ 3 at the leading edge in OPN Sh may be a continuous focal adhesion assembly which causes the cell to be attached to the substrate in a manner which would cause reduced migration. This hypothesis is confirmed by the analysis of vinculin and cortactin, which in OPN Sh showed the focal adhesions to be present around the cell periphery.

Taken together, there seems to be a prominent role of integrin $\beta$ 3 regulating a motile phenotype, either by focal adhesion assembly or formation of actin protrusions.

## **6.10 Analysis of cell-cell adhesion mediated by Slug and OPN knockdown**

### **6.10.1 Tight Junctions**

Epithelial cells are joined to each other by a set of intracellular junctions which include GAP junctions, desmosomes, adherence junctions and tight junctions. Tight junctions are formulated from occludin, claudin, and junctional adhesion molecules (JAMs). Tight junctions function to provide signalling capabilities and cell-cell adhesion aiding the maintenance of epithelial architecture. Recent reports have highlighted the tight junction down-regulation correlates with staging, invasiveness, and metastatic potential of epithelial tumours (Kominsky et al. 2003).

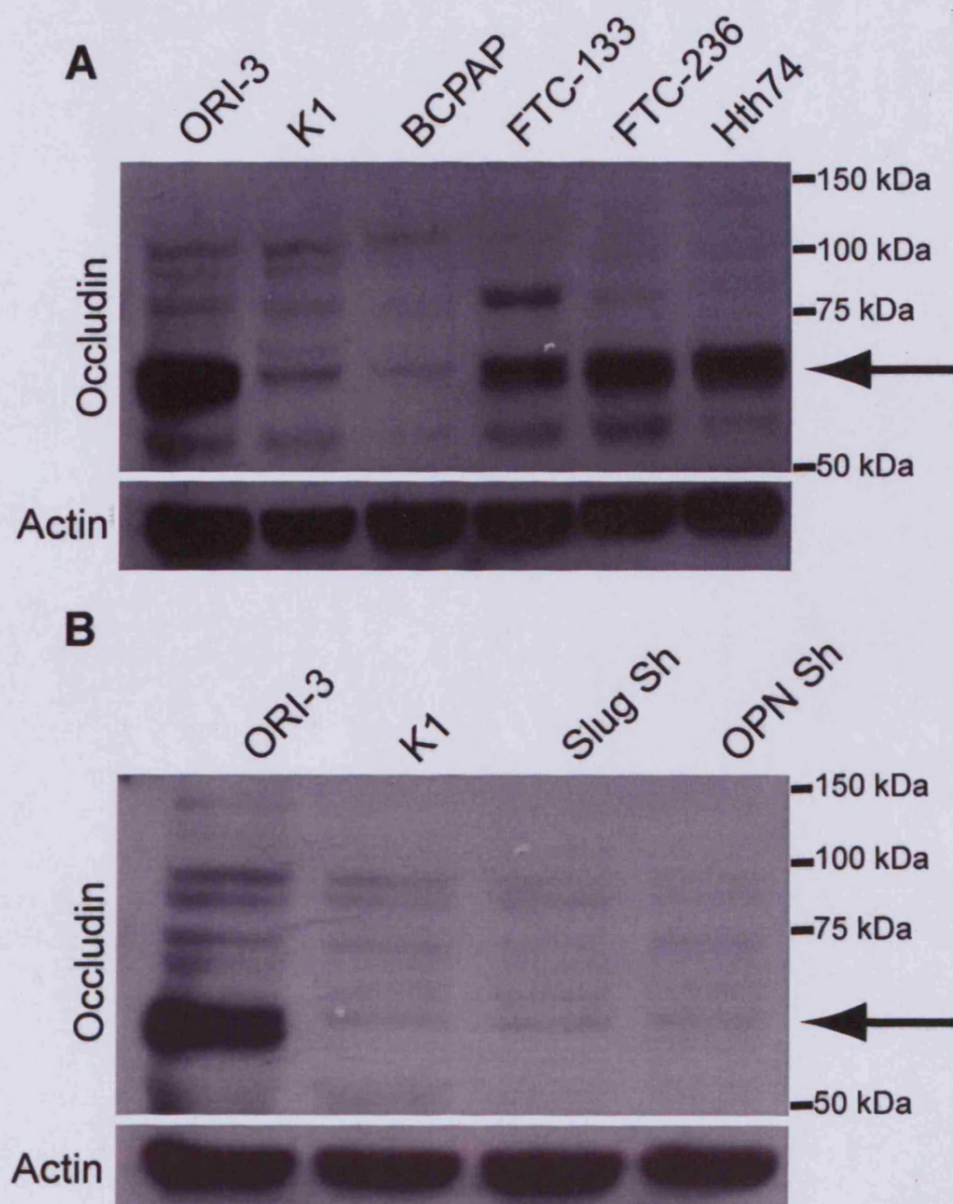
Downregulation of the tight junction protein, occludin is observed in poorly differentiated gastrointestinal adenocarcinomas (Kimura et al. 1997). This maybe important as it may be related to the progression of the cancer.

In addition, Slug is a known repressor of occludin expression in Pa4 (rat parotid gland) epithelial cells. Knockdown of Slug in these cells restored expression of occludin (Wang et al. 2007).

I sought to investigate if Slug mediates expression of occludin expression in thyroid cancer. In addition, I also investigated the role of OPN on the regulation of tight junction regulation, as there is no literature pertaining to OPN with occludin expression. To investigate this aim, I determined the expression of occludin in thyroid cell lines using SDS-PAGE western analysis. Occludin expression was highlighted in the ORI-3 cell line. Expression was also detected in the FTC cell line FTC133, and its metastasis, FTC236 and anaplastic carcinoma, Hth74. In the PTC cell lines, K1 and BCPAP occludin expression was not detected (Figure 48A).



Western blot analysis of Slug Sh and OPN Sh also did not reveal occludin expression (Figure 48B).



**Figure 48.** Analysis of tight junctions by Slug Sh and OPN Sh. SDS-PAGE western blot analysis was conducted to assess occludin expression. **A**, Thyroid cell lines were analysed for occludin expression using western blot analysis **B**, Using positive (ORI-3) and negative (K1) controls for occludin expression the knockdown clones of Slug and OPN were analysed for expression. Occludin is indicated by the black arrow. Remaining bands are non-specific.

### 6.10.2 Adherence junctions

Classical cadherins are transmembrane proteins which formulate the adherence adhesive junctions. They formulate intercellular contact through the interaction with opposing cadherins on neighbouring cells. In general, cadherins are complexed with other cytoplasmic proteins to regulate intracellular signalling. Functionally, adherence junctions initiate and stabilise cell-cell junctions to regulate the actin cytoskeleton. Epithelial cells highly express E-cadherin, whilst other cadherins are expressed by mesenchymal cells. In a large number of epithelial tumours E-cadherin is downregulated, and is linked to tumour progression. Slug is a well known transcriptional repressor of E-cadherin (Hajra et al. 2002; Uchikado et al. 2005). In the knockdown of Slug I would expect re-expression of E-cadherin.

To investigate this hypothesis, I conducted RT-PCR analysis to determine E-cadherin mRNA content in Slug Sh and OPN Sh. This experiment aimed to determine if Slug or OPN regulate E-cadherin expression. I conducted RT-PCR using MCF7 and MDAMB231 as additional controls. E-cadherin mRNA was only detected in the MCF7 cell line. It was not detectable in the ORI-3 cell line, a result that is surprising since ORI-3 cells are derived from a normal epithelial origin (Figure 49A). This indicates that E-cadherin is not present in thyroid cell lines.

I next performed immunofluorescence analysis to observe the localisation of E-cadherin (Figure 50). This method was used to re-confirm the results seen with the RT-PCR analysis. This was conducted using A431 and ORI3 cells as positive controls for E-cadherin expression. E-cadherin was only detected in the A431 cell line. Expression was localised to the cell-cell contacts, which is consistent with the normal localisation of E-cadherin. In the thyroid cell lines there was no detection of

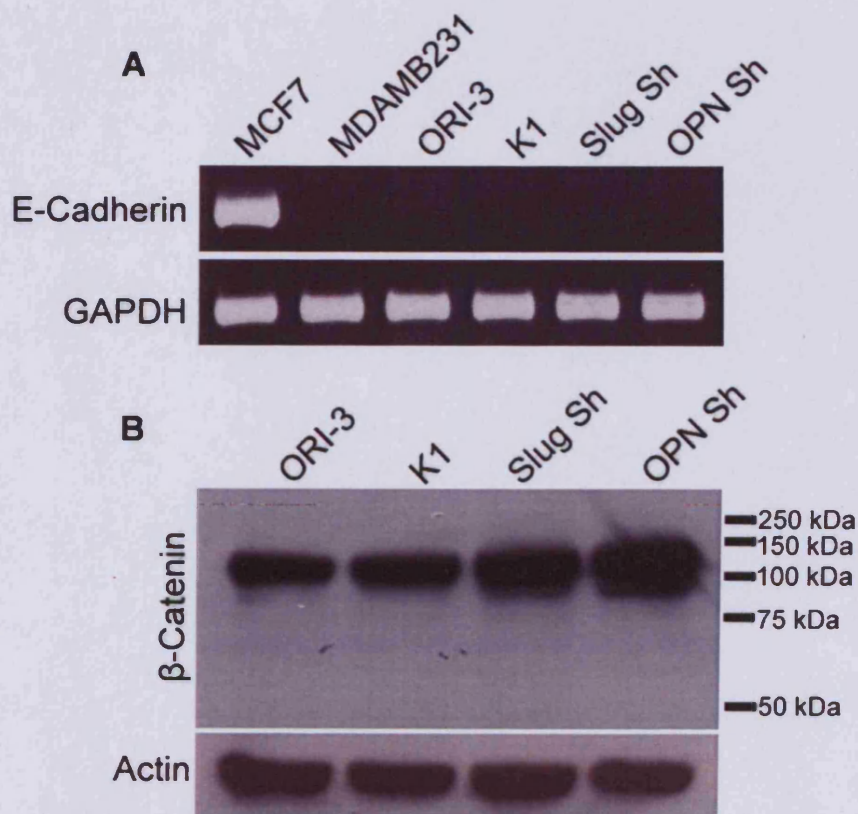
E-cadherin, a result which is consistent with the RT-PCR analysis. This result concludes further experiments regarding E-cadherin as there is no E-cadherin mRNA in thyroid.

The cytoplasmic tails of cadherin molecules bind to catenin molecules. The  $\beta$ -catenin molecule links the cadherin to the actin cytoskeleton of the cell. Indeed, the localisation of  $\beta$ -catenin is crucial to the signalling activities of the Wnt pathway.

I next investigated the expression and localisation of  $\beta$ -catenin to assess whether it is localised to the cell membrane. The overall expression of  $\beta$ -catenin was analysed using western blot analysis and to observe specific  $\beta$ -catenin localisation the technique of immunofluorescence was utilised (Figure 57B). Western blot analysis showed the expression of  $\beta$ -catenin to be increased in both Slug Sh and OPN Sh.  $\beta$ -catenin expression was also detected in the ORI-3 and K1 cell lines.

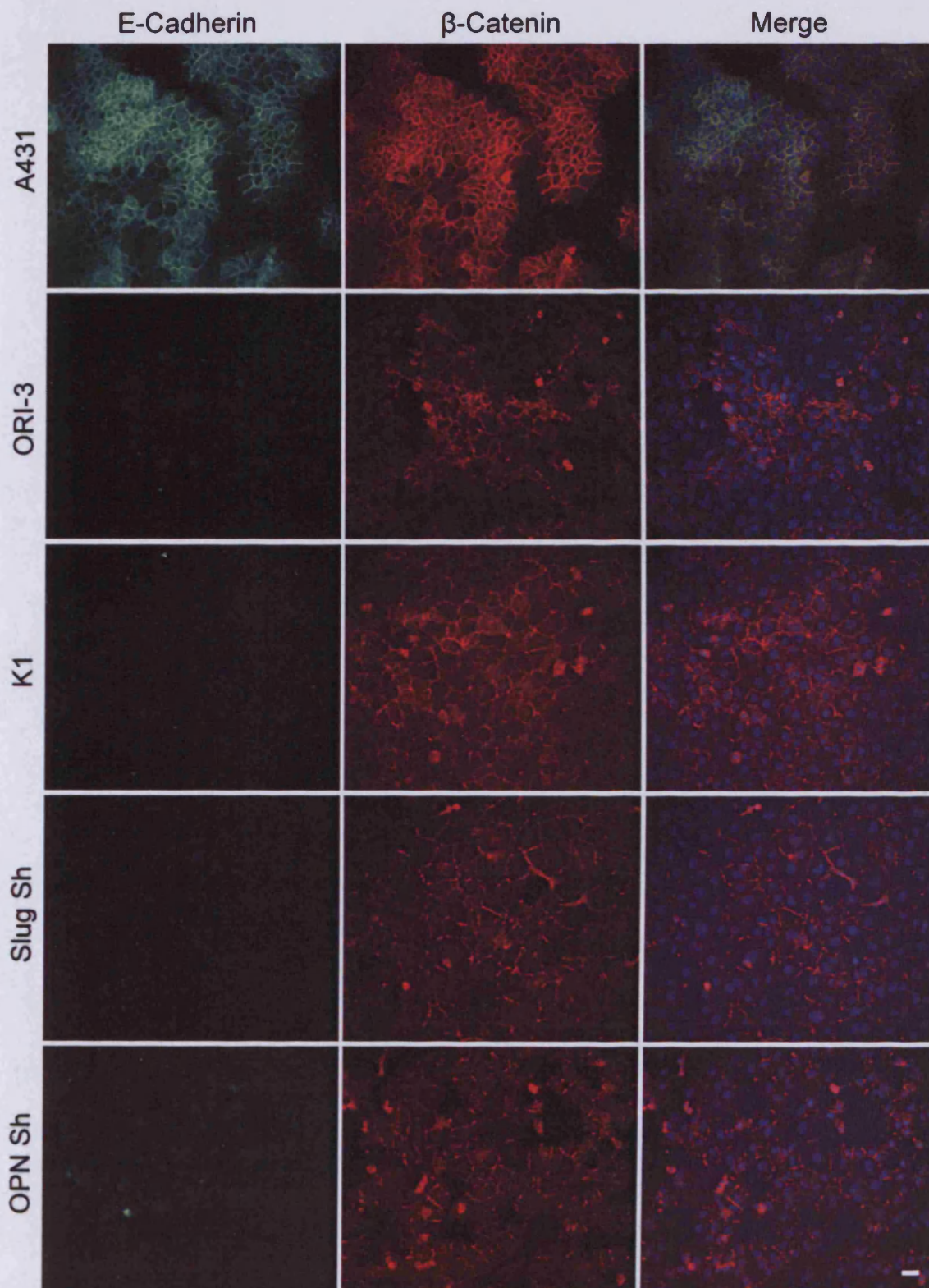
Immunofluorescence analysis showed  $\beta$ -catenin to be localised at the cell membrane. The immunofluorescence also showed that A431 positively expressed E-cadherin. This effectively showed that  $\beta$ -catenin is tethered to E-cadherin at the cell membrane. Taken together, the localisation of  $\beta$ -catenin at the cell membrane of the thyroid cell lines suggests that in the thyroid cell lines of ORI-3, K1, Slug Sh and OPN Sh  $\beta$ -catenin must be tethered to another cadherin family member (Figure 49B). To further this investigation, I utilised a pan-cadherin antibody to assess localisation of cadherin molecules. This antibody is useful in detecting all cadherin members and would show the presence and localisation of cadherins, but it cannot differentiate the type of cadherin. To investigate the localisation of cadherin molecules, sub-confluent cells plated on cover slips were fixed with paraformaldehyde and immunofluorescently stained for pan-cadherin (Figure 51).

The results showed intense pan-cadherin expression localised at the cell membrane in the ORI-3 cell line. This is typical of an epithelial cells which have adherence junctions as part of their morphology. In K1 cells, pan-cadherin was detected at the cell membrane and cytoplasm, implying a disorganised cadherin expression. This indicates the presence of a cadherin member which may mediate the invasive phenotype observed in K1 cells. The Slug Sh cells displayed positive pan-cadherin expression at the cell-cell contacts, in a highly organised structure, consistent with epithelial architecture. The OPN Sh cells displayed a disorganised pan-cadherin expression, localised at the cell membrane and the cytoplasm.

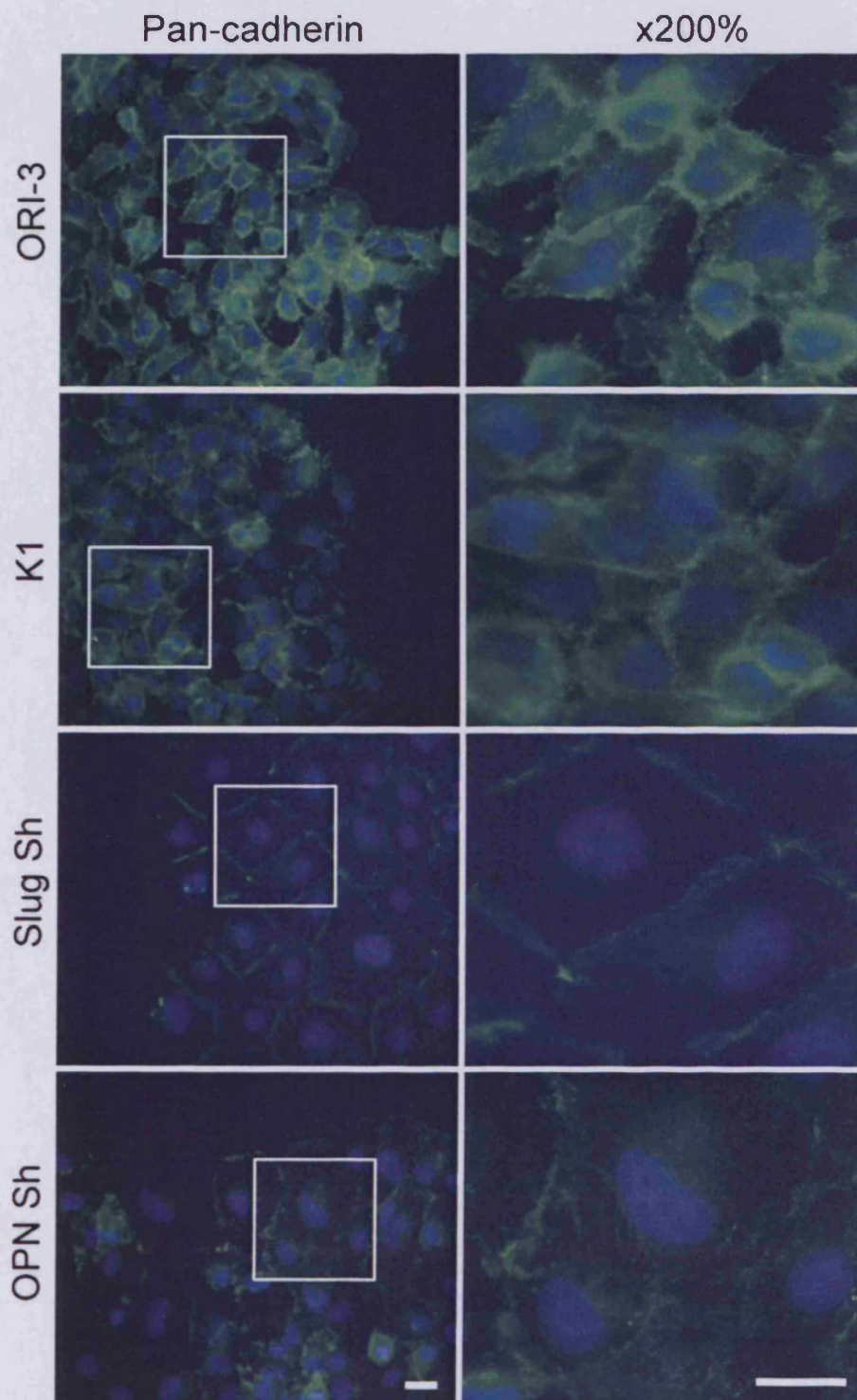


**Figure 49.** Analysis of E-cadherin expression and mRNA content. RT-PCR was conducted on subconfluent cells. Cells were lysed with Trizol and total RNA extracted, converted to cDNA using reverse transcriptase and an E-cadherin fragment amplified with the use of specific primers.  $\beta$ -catenin expression was analysed by western blot with actin acting as a loading control for the western.





**Figure 50.** Immunofluorescent analysis of E-cadherin and  $\beta$ -catenin expression. Cells were stained with E-cadherin (green) and co-immunostained with  $\beta$ -catenin (red) followed with counterstaining with DAPI nuclear stain (blue). Fluorescence was observed with a fluorescent microscope. Scale bar represents 50 $\mu$ m, and is applicable to all images.



**Figure 51.** Pan-cadherin expression in ORI-3, K1, Slug Sh and OPN Sh cells. Cells were grown to confluence on coverslips wounded for 3 hours and fixed with paraformaldehyde. Cells were stained with Pan-cadherin (green) and counterstained with DAPI nuclear stain (blue). Fluorescence was observed with a fluorescent microscope with an x40 objective lens. Images are presented as Pan-cadherin merged with DAPI for each cell line analysed. Areas were selected (signified by white squares) and enlarged by 200% to provide further intricate detail. Scale bars represents 50µm for each column.



### 6.10.3 Summary

Epithelial architecture is maintained by tightly packed cells which adhere to one another through different types of junctions, which include tight junctions, adherence junctions, GAP junctions and desmosomal junctions.

Slug functions as a transcription repressor of the adherence junctional protein E-cadherin. It also represses other junctional proteins including ZO-1 and occludin (Bolos et al. 2003; Ikenouchi et al. 2003; Ohkubo and Ozawa 2004). Overexpression of Slug has been shown to correlate with the dysregulation of occludin in a Raf-1 dependent manner (Wang et al. 2007).

Expression of occludin was absent from the K1 cell line, and in the Slug Sh and OPN Sh cells there was no alteration in occludin expression. Experimental evidence did indicate that occludin is absent from PTC. The difference in the morphology of these cells may account for this.

Although I have shown that tight junctions are not regulated as a mechanism of invasion, I analysed the possibility that adherence junctions may be down-regulated as an invasive mechanism. The loss of E-cadherin expression represents a key mechanism which occurs in EMT. Observed in thyroid cancer, Slug is a known repressor of E-cadherin expression (Hardy et al. 2007). In normal conditions, E-cadherin is associated with  $\beta$ -catenin, an integral constituent of the Wnt signalling pathway. The localisation of  $\beta$ -catenin is crucial to determining the cellular fate and transcription of genes regulating which may have roles in regulating an invasive phenotype, namely Slug. E-cadherin mRNA content was analysed in a number of cell lines with the adenocarcinoma cell line, MCF7 acting as an experimental control. In contrast to other studies, ORI-3 cells did not have any E-cadherin mRNA content (Hardy et al. 2007). A probable explanation causing the loss of E-cadherin

expression may involve the extensive culturing. Recently investigated in prostate cancer, OPN was shown to mediate an up regulation of  $\beta$ -catenin expression in the nucleus in an Akt-dependent manner (Robertson and Chellaiah).

In OPN Sh and Slug Sh cells, the western blot analysis revealed an increase in  $\beta$ -catenin expression. Immunofluorescent analysis revealed increased organised expression at the cell-cell borders in Slug Sh and OPN Sh, with minimal expression in the nucleus.

The localisation of  $\beta$ -catenin at the cell membrane indicates the presence of a cadherin molecule mediating the adherence junction between cells. Obviously this cannot be E-cadherin since this was devoid in the thyroid cell lines. The use of a pan-cadherin antibody was useful in investigating if there is in fact a cadherin present at the cell membrane. From the immunofluorescence analysis, there was intense staining at the cell membrane of ORI-3 and, somewhat K1 cells. Perhaps most interesting is the increased organised nature of the cadherin staining at the cell membrane for Slug Sh, a morphology consistent with normal epithelial cells. It appears down-regulating Slug confers an organised epithelial structure. In OPN Sh, the localisation of pan-cadherin was chaotic, and indicates that OPN does not regulate the adherence junctions as an invasive mechanism.

To summarise, the highly organised nature of the cell-cell-contacts identified by the intense  $\beta$ -catenin and pan-cadherin staining in the knockdown of Slug which implies that there is restoration of the cell-cell contacts which are evident in normal epithelium.

Cell-cell contact is crucial to maintaining an epithelial architecture, and this forms the basis of explaining how Slug mediates an invasive phenotype.





# **7.0 General Discussion and Future Work**

Thyroid cancer is the most common endocrine malignancy, with its incidence increasing worldwide. Fine needle aspiration (FNA) is the preferred preoperative diagnostic tool which can diagnose papillary carcinoma (PTC) with high reliability. FNA cannot differentiate follicular adenoma (FA) from FTC as these lesions have overlapping pathological features. The diagnosis relies upon histological examination and is an invasive technique.

This thesis contained 2 major aims, firstly, to identify novel markers of tumour invasion, and secondly to characterise their therapeutic value. To conduct this research my group had previously established *in vitro* models of follicular adenoma (FA) and papillary carcinoma (PTC). The respective oncogenes, RASV12, RET/PTC1, BRAF<sup>V600E</sup> were introduced into primary thyrocytes (Bond et al. 1994). These formed the *in vitro* models of FA (RASV12), and PTC (RET/PTC1 and BRAF<sup>V600E</sup>). Studies concerning these particular models and invasion remained a novel aspect which had not been previously investigated.

For the first major aim of my thesis, I primarily focused my investigation on the BRAF<sup>V600E</sup> oncogene and characterising the B-Raf oncogenic mutation in thyroid cell lines. B-Raf can undergo alternative splicing providing different levels of signalling which may lead to a specific cellular effect. Mutated B-Raf conveys a dramatic increase in kinase activity and is implicated in aggressive PTC tumours (Namba et al. 2003; Palona et al. 2006). The detection of BRAF<sup>V600E</sup> in both the full length and shorter isoforms in thyroid cell lines was important for future experiments and highlights the complexity of oncogenic signalling events. All the PTC cell lines which were analysed had the BRAF<sup>V600E</sup> mutation. It would be interesting to firstly investigate the kinase activity of the shorter isoforms, and the contribution to the

oncogenic ability and secondly, compare the kinase activity of the full length BRAF<sup>V600E</sup> which I used in my experiments with the other shorter isoforms.

Following on from this work, I formed my *in vitro* models for use on a microarray platform. Since I aimed to identify candidate genes involved with tumour invasion and define their therapeutic value, I characterised the invasive capability of the *in vitro* models in parallel to the genetic profiling for the identification of novel markers.

The FA model displayed a high degree of invasive behaviour, unexpected since RASV12 is associated with FA tumours. This may in part be due to culture conditions, where the serum may contain growth factors that induce invasive behaviour. The other PTC models also displayed an invasive phenotype.

This therefore limited the initial purpose of the microarray study. Since all the models displayed an invasive phenotype, the comparison of an invasive model versus a non-invasive model was not possible. The gene profiling study was reframed to determine the overlapping expression associated with or attributed to invasion for each of the models. The further aim was to identify specific processes, pathways and genes which may cause the invasive behaviour of the RASV12 model. A number of genes were identified which other studies suggest may convey the invasive phenotype seen in the RASV12 model. These genes included VEGF, FN, EGFR, MMP-9 and Rho genes. The gene expression profiles of each model showed that these genes were not up-regulated in the PTC models. The oncogenic events of BRAF<sup>V600E</sup>, RASV12 and RET/PTC1 are mutually exclusive. Therefore the genetic profile, processes and pathways are specific to the model. These genes were only analysed using the microarray platform and so to be accurate with any profiling study, it is necessary to validate the results to observe expression at the

transcriptional level (Northern blot) and protein level (SDS-PAGE) is essential. To investigate the role of serum on inducing invasive behaviour, culturing the cells in low serum conditions may result in a reduction of invasive ability in RASV12. Using the same conditions for the microarray study genetic profiling of models cultured in low serum conditions would provide a comparison of how low serum culture conditions affect gene expression. Ultimately, adjusting the culture conditions may provide a more suitable model to investigate the concept of thyroid tumour invasion.

An alternative explanation of why the RASV12 model displayed an invasion phenotype are the high proliferative sub-culture of epithelial like cells (Bond et al. 1993). Investigating the phenotypic heterogeneity and sustainability in culture needs to be carried out. The isolation of the subculture and RASV12-infected epithelial cells and study of their invasive behaviour using invasion assays would determine if the invasion observed in my studies is actually RASV12-infected epithelial cells. Also, mapping cells with epithelial or mesenchymal markers before and after conducting invasion assays would provide further experimental evidence of which cells are invading. However, this may prove challenging as cells undergo complex morphogenic changes during Matrigel cell invasion which may make identifying cellular markers difficult.

The invasion assay itself was conducted using growth factor reduced Matrigel. Thyroid invasion is defined as cell invasion through the thyroid capsule. The capsule is comprised of a mixture of components including collagen and fibrous tissue. To represent this Matrigel invasion assays could be adapted to incorporate these elements to make the invasion assay experimental technique more accurate to the *in vivo* setting.

From the microarray study and background literature analysis a number of candidates were identified. This formed the second main aim of the thesis. Osteopontin (OPN) and Slug were selected as they were both implicated in tumour progression (Cano et al. 2000; Hajra et al. 2002; Wai and Kuo 2004; Hardy et al. 2007), and investigations of these genes in thyroid cancers remained limited and could be expanded upon. OPN is over-expressed and enhances papillary carcinoma invasiveness in RET/PTC transformed rat thyroid cells (Guarino et al. 2005). In thyroid tumours, OPN expression correlates with aggressive clinicopathological features of papillary carcinoma, presence of lymph node metastases and tumour size (Guarino et al. 2005). Mechanisms of how increased OPN gene expression enhances invasion and which invasive processes and pathways are involved in thyroid cancer remains unclear. Slug is a known transcriptional repressor of E-cadherin (Hajra et al. 2002) and induces epithelial to mesenchymal transition (EMT) (Savagner et al. 1997). In thyroid, Hardy et al. showed increased detection of Slug in PTC cell lines, with basal level expression in normal thyroid tissue or normal tissue derived cell lines. Mechanistic evidence of the invasive processes mediated by Slug in thyroid cancer can also be investigated further.

To investigate Slug and OPN in thyroid cancer invasion I began by characterising a panel of thyroid cancer cell lines for their invasive capability and expression of Slug and OPN. The expression of both Slug and OPN correlated with enhanced invasion of the PTC cell line, K1. Thus, indicating of roles for Slug and OPN in thyroid cancer.

To investigate the potential role of Slug and OPN in tumour invasion, shRNA-mediated knockdown was conducted using the K1 cell line. Producing the knockdown clones and re-analysing the invasive behaviour revealed a dramatic

reduction in invasive behaviour. This provided direct evidence that correlated Slug and OPN expression with tumour invasion. This evidence was consistent with other studies where different cancers displayed increased expression of Slug or OPN which correlated with malignancy (Wai and Kuo 2004; Shih et al. 2005).

Details regarding the mechanisms of how Slug and OPN regulate invasion remain to be elucidated. Since invasion consists of multiple processes which combine and initiate invasion, a multi-step analysis focussing on proteolysis, cell-cell adhesion, cell motility and cell-matrix adhesion were conducted.

Proteinases degrade ECM components to promote angiogenesis and promote invasion. I sought to analyse if proteinase activity is regulated by either Slug or OPN. Expression of MT1-MMP mRNA was reduced in both the Slug and OPN knockdown clones. This indicates that both OPN and Slug may have a role in the regulation of MT1-MMP. This membrane bound MMP promotes tumour growth and angiogenesis through the induction of vascular endothelial growth factor (Sounni et al. 2002). In addition, MMP-2 activation may be impaired in Slug and OPN knockdown clones since activation of MMP-2 requires 2 molecules of MT1-MMP. The first MT1-MMP molecule binds to the latent MMP-2 and the second cleaves and activates the latent enzyme.

The zymogram analysis revealed latent pro-enzymes of MMP-2 and MMP-9; however, it is probable that the concentration of MMP-2 and MMP-9 in the samples was not efficient to detect the active enzymes. In addition, analysis by western blot or ELISA using concentrated conditioned media could identify the pro- and active forms of MMP-2 and MMP-9. This would confirm the results of the zymogram.

MT1-MMP was shown to collaborate with integrin $\beta$ 3 to promote MMP-2 activation in breast carcinoma cells (Deryugina et al. 2001). Analysing the expression of integrin $\beta$ 3 would elucidate what the effect of MT1-MMP down-regulation is in the knock-down clones. Since MT1-MMP is membrane bound, using immunofluorescence to analyse localisation, it would be interesting to observe if integrin $\beta$ 3 co-localises with MT1-MMP.

The urokinase activation system did not have a role in invasion mediated by Slug or OPN since decrease in mRNA content of uPAR was not seen. Further investigation would involve observing expression of uPAR and uPA using western blots to observe expression at the protein level. In addition, detection of uPA by RT-PCR does not differentiate between the active and inactive forms of uPA in the samples. To confirm this evidence presented, an assay testing for uPA activity would need to be conducted.

The 2nd aspect of investigation was the investigation of cell-cell interactions. Epithelial cells are joined to each other by a set of intracellular junctions including adherence and tight junctions to give the typical epithelial architecture. In a tumour setting these junctions are down-regulated, and Slug is a main instigator of down-regulating cell junctions (Hajra et al. 2002; Bolos et al. 2003). E-cadherin is found expressed at epithelial adherence junctions, and down-regulation is associated with EMT. My investigation of E-cadherin expression was in contrast to Hardy et al. The ORI-3 cell line, derived from normal epithelium was used as a positive control for E-cadherin expression. In my hands E-cadherin was absent in this cell line, and this was true of all the thyroid cell lines. However, this down-regulation of E-cadherin did not affect the invasive behaviour. Hardy et al. found the same ORI-3 cell line to



possess E-cadherin. Continuous culture of the ORI-3 cell line may have resulted in the loss of E-cadherin.

Using a pan-cadherin antibody which detects all cadherin molecules, results showed that there is expression of cadherin molecules at the cell junctions, and this was confirmed with the analysis of  $\beta$ -catenin which was also localised at the cell junctions.

Analysing the knock-down clone of Slug using the pan-cadherin antibody revealed a highly organised localisation of cadherin at the cell membrane. This type of localisation is consistent with a normal epithelial monolayer. To investigate this further, determining which class of cadherin is expressed would be advantageous. Class switching of cadherins may be a mechanism which may be regulated by Slug (Kuphal and Bosserhoff 2006).

The third aspect of this investigation observed cell motility. In a tumour environment, cells possess a motile phenotype in order to invade. Motility is dependent on the formation of actin protrusions (lamellipodia and filopodia) and cell-matrix adhesive contacts which all combine to generate forward thrust (Insall and Machesky 2009).

Time-lapse analysis tracking cells and conducting random walk analysis revealed phenotypic changes resulting from the knock-down of Slug or OPN. In Slug Sh and OPN Sh cells, proliferation and the distance travelled were both decreased. In addition, the actin time course wound closure experiment revealed a slower wound closure time for the knockdown clones. Random walk analysis revealed decreased migration of Slug, however more so of the OPN knockdown clone. This evidence provided an implication that cell migration is impaired when Slug or OPN are

knocked down. These results clearly showed an impaired motile ability in the knockdown clones.

The formation of actin-based protrusions are important for cell migration, and N-WASP and WAVE1 are important to this. Immunofluorescence analysis of N-WASP in both the Slug and OPN knockdown clones revealed decreased N-WASP signal, an indication of decreased N-WASP activity. This work would greatly benefit from western blot analysis to confirm the indication. Detection of WAVE1 also revealed decreased expression in the Slug knockdown. This evidence indicates the activation of the Arp2/3 complex may be impaired, affecting the formation of lamellipodia and filopodia. N-WASP is activated by cdc42 (Wu et al. 2004), and Rac functions upstream of WAVE1 (Schwartz 2004). Both cdc42 and Rac function at the leading edge of the cell whereas Rho acts at the rear of the cell (Popoff and Geny 2009). GST-fusion assays to examine Rac and Rho activity were conducted using the Rac binding domain of PAK1 and the Rho binding domain of Rhotekin. However, the assay requires optimisation and so did not provide any meaningful results. A GST-pull down assay testing for cdc42 activity would also be advantageous to see if these GTPases are regulated by Slug or OPN.

The Rho GTPase regulates movement at the rear of the cell by forming stress fibres (Maekawa et al. 1999). The OPN knock-down clone displayed a lack of stress fibres and in a wound closure environment the cell body was stretched implying the movement at the rear of the cell is impaired. Analysis of Rho activity would elucidate if this is the case, and if OPN regulates Rho activity.

Fundamental for cell movement are the formation of focal adhesions, the size and localisation of which is dependent on the migratory status of the cell. Immunofluorescence analysis of vinculin, a protein present in mature focal

adhesions revealed cell-matrix interactions in the knockdown clones. Phosphorylated-cortactin was co-stained, which is particularly important in cells with a migratory phenotype. Cortactin is an actin filament binding protein localised at the leading edge of lamellipodia. Cortactin and N-WASP synergistically cooperate to stimulate Arp2/3 and actin assembly. When cortactin is phosphorylated, activation of N-WASP is enhanced. In the K1 PTC cell line, cortactin was localised around the cell periphery and Slug knockdown displayed a similar pattern. Although phosphorylated-cortactin was detectable in the OPN knockdown there was decreased expression of cortactin around the cell periphery.

There is circumstantial evidence of increased size of focal adhesions in the knockdown clones. This evidence links the cell-matrix adhesions with a migratory phenotype where cells with decreased vinculin expression have fewer cell-matrix adhesive structures with an increased migratory ability (Subauste et al. 2004). The partial restoration of focal adhesions in the knockdown clones for Slug and OPN links focal adhesion size with cell migration and therefore invasion. To further this work other proteins involved with cell-matrix adhesions would determine the type of adhesive contacts present. The integral focal adhesion adaptor protein paxillin has many binding partners, and performs a central role in the formation of focal contacts. Since focal adhesions remain an active site for actin polymerisation, zyxin, an adaptor protein which serves as a scaffold protein at focal adhesions has been shown to regulate actin polymerisation, and is a protein which is only present in mature focal adhesions (Beckerle 1997).

FAK is a tyrosine kinase which is present only in focal adhesions (Brakebusch and Fassler 2003). The carboxyl terminus of FAK contains the focal adhesion targeting

(FAT) domain, which displays homology to vinculin (Arold et al. 2002), and contains binding sites for paxillin and talin. Integrin-linked kinase (ILK) is a non-receptor serine/threonine kinase with a principle role as an adaptor protein. The functionality of ILK has been highlighted by ILK-null fibroblasts which exhibit reduced cell spreading, migration, due to delayed formation of focal adhesions, and a poorly organised actin cytoskeleton (Sakai et al. 2003). Investigating the expression and localisation of paxillin, zyxin, ILK and FAK would elucidate the role of focal adhesions with actin assembly/disassembly inducing cell contractility and subsequent cell migration. Formation of adhesive structures is a dynamic process during cell migration, and so observing localisation and expression of other associated proteins may provide additional mechanistic detail explaining how OPN and Slug affect the formation of cell-matrix adhesions.

Other experiments to consider would be to analyse the upstream inducers of Slug and OPN. Highlighted in the background chapter, HGF-dependent tumourigenesis and inhibition of OPN expression attenuates the HGF-invasiveness in vitro and in vivo (Medico et al. 2001; Ariztia et al. 2003). This is an indication that OPN is a necessity for HGF-induced tumour invasion. Using time-lapse analysis I found that both Slug and OPN knockdown clones did not respond to HGF treatment in terms of movement. Interestingly, in the Slug Sh and OPN Sh cells the time between cell division was increased in response to HGF treatment. This indicates that HGF may have a role in Slug and OPN-mediated invasion. To expand on this further I would conduct invasion assays with and without HGF treatment. This would therefore confirm the action of HGF in mediating invasion via Slug and OPN. Within the literature Slug is reported to be highly up-regulated in response to HGF, and this is

an early event preceding EMT (Leroy and Mostov 2007). The authors suggest Slug may act as a survival factor in response to HGF.

Variant isoforms of CD44 (CD44v) are able to amplify HGF-mediated invasion (Smith et al. 1999) and are able to bind OPN. The interaction of CD44v with integrin $\beta$ 1 can lead to enhanced cell motility and chemotaxis (Katagiri et al. 1999). I therefore would like to assess the localisation of integrin $\beta$ 1 and see if co-localisation with CD44v occurs. With regard to CD44v analysis of expression and localisation of CD44v would confirm the findings presented and elucidate the role of HGF in Slug and OPN-mediated invasion. There is a good indication that CD44v have a role since CD44v6, v7 and v10 mRNA was down-regulated by both Slug and OPN in the knockdown clones.

Craig et al., showed that interaction between Slug and CD44 induced cell invasion. This is the only study to correlate CD44 and Slug with tumour invasion. This interaction depends solely on variant isoforms of CD44, a finding consistent with my experiments. The down-regulation of Slug expression correlates with the reduced mRNA content of CD44v6, v7 and v10. This interaction can be investigated in more detail. It would be interesting to observe if the introduction of CD44v in Slug Sh or OPN Sh induces a restoration of invasive ability.

An interesting investigation would be to investigate the upstream events of both Slug and OPN. The over-expression of OPN has been correlated with the presence of BRAF<sup>V600E</sup> and RET/PTC1 (Guarino et al. 2005). It would be interesting to see if the B-Raf oncogene regulates the expression of Slug and OPN to induce invasion.

Since this knockdown work has had a phenotypic response in a cell line model, carrying this work forward with primary models of PTC would give an indication of

the role Slug and OPN in an *in vivo* situation. The evidence presented in this thesis implies that both Slug and OPN respectively are sufficient alone to initiate tumour invasion. It is unclear how knockdown of either Slug or OPN would affect the gross morphology of an *in vivo* tumour. Since growth rates are reduced along with cell migration, it would imply reduced sized tumours with a decreased capacity to invade. This would not be an absolute ablation of invasion since residual invasion was observed in the experiments conducted with the K1 cell line. However it suggests that these may provide valid candidates as therapeutic targets since in this thesis direct targeting of Slug and OPN alone severely reduced invasion, implying that these genes have a master role in mediating invasion. In other cancers targeting these genes *in vivo* situation reduced tumour invasion and tumour size (Yang et al. ; Zhao et al. 2008). The research presented in other studies together with the evidence presented in this thesis strongly indicates that targeting either Slug or OPN *in vivo* would have a strong therapeutic effect.

# **8.0 References**

Aberle, H., Bauer, A., Stappert, J., Kispert, A., and Kemler, R. 1997. beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J* 16(13): 3797-3804.

Alberga, A., Boulay, J.L., Kempe, E., Dennefeld, C., and Haenlin, M. 1991. The snail gene required for mesoderm formation in *Drosophila* is expressed dynamically in derivatives of all three germ layers. *Development* 111(4): 983-992.

Al-Brahim, N. and Asa, S.L. 2006. Papillary thyroid carcinoma: an overview. *Arch Pathol Lab Med* 130(7): 1057-1062.

Aldred, M.A., Huang, Y., Liyanarachchi, S., Pellegata, N.S., Gimm, O., Jhiang, S., Davuluri, R.V., de la Chapelle, A., and Eng, C. 2004. Papillary and follicular thyroid carcinomas show distinctly different microarray expression profiles and can be distinguished by a minimum of five genes. *J Clin Oncol* 22(17): 3531-3539.

Aldred, M.A., Huang, Y., Liyanarachchi, S., Pellegata, N.S., Gimm, O., Jhiang, S., Davuluri, R.V., de la Chapelle, A., and Eng, C. 2004. Papillary and follicular thyroid carcinomas show distinctly different microarray expression profiles and can be distinguished by a minimum of five genes. *J Clin Oncol* 22(17): 3531-3539.

Al-Jamal, R. and Harrison, D.J. 2008. Beta1 integrin in tissue remodelling and repair: from phenomena to concepts. *Pharmacol Ther* 120(2): 81-101.

Andreasen, P.A., Egelund, R., and Petersen, H.H. 2000. The plasminogen activation system in tumor growth, invasion, and metastasis. *Cell Mol Life Sci* 57(1): 25-40.

Arighi, E., Borrello, M.G., and Sartiola, H. 2005. RET tyrosine kinase signaling in development and cancer. *Cytokine Growth Factor Rev* 16(4-5): 441-467.

Arold, S.T., Hoellerer, M.K., and Noble, M.E. 2002. The structural basis of localization and signaling by the focal adhesion targeting domain. *Structure* 10(3): 319-327.

Ariztia, E.V., Subbarao, V., Solt, D.B., Rademaker, A.W., Iyer, A.P., and Oltvai, Z.N. 2003. Osteopontin contributes to hepatocyte growth factor-induced tumor growth and metastasis formation. *Exp Cell Res* 288(2): 257-267.

Attwell, S., Mills, J., Troussard, A., Wu, C., and Dedhar, S. 2003. Integration of cell attachment, cytoskeletal localization, and signaling by integrin-linked kinase (ILK), CH-ILKBP, and the tumor suppressor PTEN. *Mol Biol Cell* 14(12): 4813-4825.

Bader, A.G. and Vogt, P.K. 2004. An essential role for protein synthesis in oncogenic cellular transformation. *Oncogene* 23(18): 3145-3150.

Baitei, E.Y., Zou, M., Al-Mohanna, F., Collison, K., Alzahrani, A.S., Farid, N.R., Meyer, B., and Shi, Y. 2009. Aberrant BRAF splicing as an alternative mechanism for oncogenic B-Raf activation in thyroid carcinoma. *J Pathol* 217(5): 707-715.



- Bajorath, J., Greenfield, B., Munro, S.B., Day, A.J., and Aruffo, A. 1998. Identification of CD44 residues important for hyaluronan binding and delineation of the binding site. *J Biol Chem* 273(1): 338-343.
- Baloch, Z.W. and Livolsi, V.A. 2002. Follicular-patterned lesions of the thyroid: the bane of the pathologist. *Am J Clin Pathol* 117(1): 143-150.
- Barbera, M.J., Puig, I., Dominguez, D., Julien-Grille, S., Guaita-Esteruelas, S., Peiro, S., Baulida, J., Franci, C., Dedhar, S., Larue, L. et al. 2004. Regulation of Snail transcription during epithelial to mesenchymal transition of tumor cells. *Oncogene* 23(44): 7345-7354.
- Barden, C.B., Shister, K.W., Zhu, B., Guiter, G., Greenblatt, D.Y., Zeiger, M.A., and Fahey, T.J., 3rd. 2003. Classification of follicular thyroid tumors by molecular signature: results of gene profiling. *Clin Cancer Res* 9(5): 1792-1800.
- Baris, O., Mirebeau-Prunier, D., Savagner, F., Rodien, P., Ballester, B., Lloriod, B., Granjeaud, S., Guyetant, S., Franc, B., Houlgatte, R. et al. 2005. Gene profiling reveals specific oncogenic mechanisms and signaling pathways in oncocytic and papillary thyroid carcinoma. *Oncogene* 24(25): 4155-4161.
- Barnier, J.V., Papin, C., Eyche, A., Lecoq, O., and Calothy, G. 1995. The mouse B-raf gene encodes multiple protein isoforms with tissue-specific expression. *J Biol Chem* 270(40): 23381-23389.
- Bautista, D.S., Denstedt, J., Chambers, A.F., and Harris, J.F. 1996. Low-molecular-weight variants of osteopontin generated by serine proteinases in urine of patients with kidney stones. *J Cell Biochem* 61(3): 402-409.
- Beckerle, M.C. 1997. Zyxin: zinc fingers at sites of cell adhesion. *Bioessays* 19(11): 949-957.
- Bergers, G., Brekken, R., McMahon, G., Vu, T.H., Itoh, T., Tamaki, K., Tanzawa, K., Thorpe, P., Itohara, S., Werb, Z. et al. 2000. Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nat Cell Biol* 2(10): 737-744.
- Birchmeier, C. and Gherardi, E. 1998. Developmental roles of HGF/SF and its receptor, the c-Met tyrosine kinase. *Trends Cell Biol* 8(10): 404-410.
- Birgersdotter, A., R. Sandberg, et al. 2005. Gene expression perturbation in vitro--a growing case for three-dimensional (3D) culture systems. *Semin Cancer Biol* 15(5): 405-412.
- Bjornland, K., J. O. Winberg, 1999. S100A4 involvement in metastasis: deregulation of matrix metalloproteinases and tissue inhibitors of matrix metalloproteinases in osteosarcoma cells transfected with an anti-S100A4 ribozyme. *Cancer Res* 59(18): 4702-4708.

Blanco, M.J., Moreno-Bueno, G., Sarrio, D., Locascio, A., Cano, A., Palacios, J., and Nieto, M.A. 2002. Correlation of Snail expression with histological grade and lymph node status in breast carcinomas. *Oncogene* 21(20): 3241-3246.

Blobel, G.A. and Hanafusa, H. 1991. The v-src inducible gene 9E3/pCEF4 is regulated by both its promoter upstream sequence and its 3' untranslated region. *Proc Natl Acad Sci U S A* 88(4): 1162-1166.

Boettner, B. and Van Aelst, L. 2009. Control of cell adhesion dynamics by Rap1 signaling. *Curr Opin Cell Biol* 21(5): 684-693.

Bolos, V., Peinado, H., Perez-Moreno, M.A., Fraga, M.F., Esteller, M., and Cano, A. 2003. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci* 116(Pt 3): 499-511.

Bolos, V., Peinado, H., Perez-Moreno, M.A., Fraga, M.F., Esteller, M., and Cano, A. 2003. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci* 116(Pt 3): 499-511.

Bond, J.A., Wyllie, F.S., Ivan, M., Dawson, T., and Wynford-Thomas, D. 1993. A variant epithelial sub-population in normal thyroid with high proliferative capacity in vitro. *Mol Cell Endocrinol* 93(2): 175-183.

Bond, J.A., Wyllie, F.S., Rowson, J., Radulescu, A., and Wynford-Thomas, D. 1994. In vitro reconstruction of tumour initiation in a human epithelium. *Oncogene* 9(1): 281-290.

Bongarzone, I., Butti, M.G., Coronelli, S., Borrello, M.G., Santoro, M., Mondellini, P., Pilotti, S., Fusco, A., Della Porta, G., and Pierotti, M.A. 1994. Frequent activation of ret protooncogene by fusion with a new activating gene in papillary thyroid carcinomas. *Cancer Res* 54(11): 2979-2985.

Boulay, J.L., Dennefeld, C., and Alberga, A. 1987. The Drosophila developmental gene snail encodes a protein with nucleic acid binding fingers. *Nature* 330(6146): 395-398.

Brakebusch, C. and Fassler, R. 2003. The integrin-actin connection, an eternal love affair. *EMBO J* 22(10): 2324-2333.

Buergy, D., Weber, T., Maurer, G.D., Mudduluru, G., Medved, F., Leupold, J.H., Brauckhoff, M., Post, S., Dralle, H., and Allgayer, H. 2009. Urokinase receptor, MMP-1 and MMP-9 are markers to differentiate prognosis, adenoma and carcinoma in thyroid malignancies. *Int J Cancer* 125(4): 894-901.

Burridge, K. and Wennerberg, K. 2004. Rho and Rac take center stage. *Cell* 116(2): 167-179.

Calogeraki, I., Barnier, J.V., Eyche, A., Felder, M.P., Calothy, G., and Marx, M. 1993. Genomic organization and nucleotide sequence of the coding region of the chicken c-Rml(B-raf-1) proto-oncogene. *Biochem Biophys Res Commun* 193(3): 1324-1331.

Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol* 2(2): 76-83.

Carrier, M.F., Laurent, V., Santolini, J., Melki, R., Didry, D., Xia, G.X., Hong, Y., Chua, N.H., and Pantaloni, D. 1997. Actin depolymerizing factor (ADF/cofilin) enhances the rate of filament turnover: implication in actin-based motility. *J Cell Biol* 136(6): 1307-1322.

Castellone, M.D., Celetti, A., Guarino, V., Cirafici, A.M., Basolo, F., Giannini, R., Medico, E., Kruhoffer, M., Orntoft, T.F., Curcio, F. et al. 2004. Autocrine stimulation by osteopontin plays a pivotal role in the expression of the mitogenic and invasive phenotype of RET/PTC-transformed thyroid cells. *Oncogene* 23(12): 2188-2196.

Celetti, A., Cerrato, A., Merolla, F., Vitagliano, D., Vecchio, G., and Grieco, M. 2004. H4(D10S170), a gene frequently rearranged with RET in papillary thyroid carcinomas: functional characterization. *Oncogene* 23(1): 109-121.

Cerutti, J.M., Oler, G., Michaluart, P., Jr., Delcelo, R., Beaty, R.M., Shoemaker, J., and Riggins, G.J. 2007. Molecular profiling of matched samples identifies biomarkers of papillary thyroid carcinoma lymph node metastasis. *Cancer Res* 67(16): 7885-7892.

Chambers, A.F., Groom, A.C., and MacDonald, I.C. 2002. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* 2(8): 563-572.

Chen, Y.J., Wei, Y.Y., Chen, H.T., Fong, Y.C., Hsu, C.J., Tsai, C.H., Hsu, H.C., Liu, S.H., and Tang, C.H. 2009. Osteopontin increases migration and MMP-9 up-regulation via  $\alpha$ 5 $\beta$ 3 integrin, FAK, ERK, and NF- $\kappa$ B-dependent pathway in human chondrosarcoma cells. *J Cell Physiol* 221(1): 98-108.

Cherfils, J. and Chardin, P. 1999. GEFs: structural basis for their activation of small GTP-binding proteins. *Trends Biochem Sci* 24(8): 306-311.

Chevillard, S., Ugolin, N., Vielh, P., Ory, K., Levalois, C., Elliott, D., Clayman, G.L., and El-Naggar, A.K. 2004. Gene expression profiling of differentiated thyroid neoplasms: diagnostic and clinical implications. *Clin Cancer Res* 10(19): 6586-6597.

Chiswell, B.P., Zhang, R., Murphy, J.W., Boggon, T.J., and Calderwood, D.A. 2008. The structural basis of integrin-linked kinase-PINCH interactions. *Proc Natl Acad Sci U S A* 105(52): 20677-20682.

Cho Mar, K., Eimoto, T., Tateyama, H., Arai, Y., Fujiyoshi, Y., and Hamaguchi, M. 2006. Expression of matrix metalloproteinases in benign and malignant follicular thyroid lesions. *Histopathology* 48(3): 286-294.

Cho, H.J., Baek, K.E., Saika, S., Jeong, M.J., and Yoo, J. 2007. Snail is required for transforming growth factor-beta-induced epithelial-mesenchymal transition by activating PI3 kinase/Akt signal pathway. *Biochem Biophys Res Commun* 353(2): 337-343.

Cho, J.Y., Sagartz, J.E., Capen, C.C., Mazzaferri, E.L., and Jhiang, S.M. 1999. Early cellular abnormalities induced by RET/PTC1 oncogene in thyroid-targeted transgenic mice. *Oncogene* 18(24): 3659-3665.

Choi, J., Park, S.Y., and Joo, C.K. 2007. Transforming growth factor-beta1 represses E-cadherin production via slug expression in lens epithelial cells. *Invest Ophthalmol Vis Sci* 48(6): 2708-2718.

Codling, E. A., M. J. Plank. 2008. Random walk models in biology. *J R Soc Interface* 5(25): 813-834.

Conacci-Sorrell, M., Simcha, I., Ben-Yedidia, T., Blechman, J., Savagner, P., and Ben-Ze'ev, A. 2003. Autoregulation of E-cadherin expression by cadherin-cadherin interactions: the roles of beta-catenin signaling, Slug, and MAPK. *J Cell Biol* 163(4): 847-857.

Corvi, R., Berger, N., Balczon, R., and Romeo, G. 2000. RET/PCM-1: a novel fusion gene in papillary thyroid carcinoma. *Oncogene* 19(37): 4236-4242.

Craig, E.A., Parker, P., and Camenisch, T.D. 2009. Size-dependent regulation of Snail2 by hyaluronan: its role in cellular invasion. *Glycobiology* 19(8): 890-898.

Cukierman, E., R. Pankov. 2001. Taking cell-matrix adhesions to the third dimension. *Science* 294(5547): 1708-1712.

D'Alessandro M.P 1995. Anatomy Atlases [Online]

Available at:

<http://www.anatomyatlases.org/MicroscopicAnatomy/Section15/Plate15286.shtml>

Accessed: 30/11/2010.

Danen, E.H., van Rheenen, J., Franken, W., Huveneers, S., Sonneveld, P., Jalink, K., and Sonnenberg, A. 2005. Integrins control motile strategy through a Rho-cofilin pathway. *J Cell Biol* 169(3): 515-526.

Das, R., Mahabeleshwar, G.H., and Kundu, G.C. 2003. Osteopontin stimulates cell motility and nuclear factor kappaB-mediated secretion of urokinase type plasminogen activator through phosphatidylinositol 3-kinase/Akt signaling pathways in breast cancer cells. *J Biol Chem* 278(31): 28593-28606.

Das, R., Mahabeleshwar, G.H., and Kundu, G.C. 2004. Osteopontin induces AP-1-mediated secretion of urokinase-type plasminogen activator through c-Src-

dependent epidermal growth factor receptor transactivation in breast cancer cells. *J Biol Chem* 279(12): 11051-11064.

Dass, K., Ahmad, A., Azmi, A.S., Sarkar, S.H., and Sarkar, F.H. 2008. Evolving role of uPA/uPAR system in human cancers. *Cancer Treat Rev* 34(2): 122-136.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W. et al. 2002. Mutations of the BRAF gene in human cancer. *Nature* 417(6892): 949-954.

Davis, N.L., Gordon, M., Germann, E., Robins, R.E., and McGregor, G.I. 1991. Clinical parameters predictive of malignancy of thyroid follicular neoplasms. *Am J Surg* 161(5): 567-569.

Davis, M.A., Ireton, R.C., and Reynolds, A.B. 2003. A core function for p120-catenin in cadherin turnover. *J Cell Biol* 163(3): 525-534.

de la Torre, N.G., Buley, I., Wass, J.A., and Turner, H.E. 2006. Angiogenesis and lymphangiogenesis in thyroid proliferative lesions: relationship to type and tumour behaviour. *Endocr Relat Cancer* 13(3): 931-944.

DeLellis, R.A. 2006. Pathology and genetics of thyroid carcinoma. *J Surg Oncol* 94(8): 662-669.

Delys, L., Detours, V., Franc, B., Thomas, G., Bogdanova, T., Tronko, M., Libert, F., Dumont, J.E., and Maenhaut, C. 2007. Gene expression and the biological phenotype of papillary thyroid carcinomas. *Oncogene* 26(57): 7894-7903.

Denhardt, D.T., Mistretta, D., Chambers, A.F., Krishna, S., Porter, J.F., Raghuram, S., and Rittling, S.R. 2003. Transcriptional regulation of osteopontin and the metastatic phenotype: evidence for a Ras-activated enhancer in the human OPN promoter. *Clin Exp Metastasis* 20(1): 77-84.

Denhardt, D.T., Noda, M., O'Regan, A.W., Pavlin, D., and Berman, J.S. 2001. Osteopontin as a means to cope with environmental insults: regulation of inflammation, tissue remodeling, and cell survival. *J Clin Invest* 107(9): 1055-1061.

Deryugina, E.I. and Quigley, J.P. 2006. Matrix metalloproteinases and tumor metastasis. *Cancer Metastasis Rev* 25(1): 9-34.

Deryugina, E.I., Ratnikov, B., Monosov, E., Postnova, T.I., DiScipio, R., Smith, J.W., and Strongin, A.Y. 2001. MT1-MMP initiates activation of pro-MMP-2 and integrin  $\alpha$ v $\beta$ 3 promotes maturation of MMP-2 in breast carcinoma cells. *Exp Cell Res* 263(2): 209-223.

Desmouliere, A. 1995. Factors influencing myofibroblast differentiation during wound healing and fibrosis. *Cell Biol Int* 19(5): 471-476.

Dhillon, A.S., Meikle, S., Yazici, Z., Eulitz, M., and Kolch, W. 2002. Regulation of Raf-1 activation and signalling by dephosphorylation. *Embo J* 21(1-2): 64-71.

- Dikeman, D.A., Rivera Rosado, L.A., Horn, T.A., Alves, C.S., Konstantopoulos, K., and Yang, J.T. 2008.  $\alpha 4 \beta 1$ -Integrin regulates directionally persistent cell migration in response to shear flow stimulation. *Am J Physiol Cell Physiol* 295(1): C151-159.
- Ding, Z., Lambrechts, A., Parepally, M., and Roy, P. 2006. Silencing profilin-1 inhibits endothelial cell proliferation, migration and cord morphogenesis. *J Cell Sci* 119(Pt 19): 4127-4137.
- Donovan, S., Shannon, K.M., and Bollag, G. 2002. GTPase activating proteins: critical regulators of intracellular signaling. *Biochim Biophys Acta* 1602(1): 23-45.
- Edwards, D.C., Sanders, L.C., Bokoch, G.M., and Gill, G.N. 1999. Activation of LIM-kinase by Pak1 couples Rac/Cdc42 GTPase signalling to actin cytoskeletal dynamics. *Nat Cell Biol* 1(5): 253-259.
- Ekici, S., Ayhan, A., Kendi, S., and Ozen, H. 2002. Determination of prognosis in patients with prostate cancer treated with radical prostatectomy: prognostic value of CD44v6 score. *J Urol* 167(5): 2037-2041.
- El-Tanani, M., Barraclough, R., Wilkinson, M.C., and Rudland, P.S. 2001. Metastasis-inducing dna regulates the expression of the osteopontin gene by binding the transcription factor Tcf-4. *Cancer Res* 61(14): 5619-5629.
- El-Tanani, M., Platt-Higgins, A., Rudland, P.S., and Campbell, F.C. 2004. Ets gene PEA3 cooperates with beta-catenin-Lef-1 and c-Jun in regulation of osteopontin transcription. *J Biol Chem* 279(20): 20794-20806.
- El-Tanani, M.K., Campbell, F.C., Kurisetty, V., Jin, D., McCann, M., and Rudland, P.S. 2006. The regulation and role of osteopontin in malignant transformation and cancer. *Cytokine Growth Factor Rev* 17(6): 463-474.
- Eszlinger, M., Krohn, K., Kukulska, A., Jarzab, B., and Paschke, R. 2007. Perspectives and limitations of microarray-based gene expression profiling of thyroid tumors. *Endocr Rev* 28(3): 322-338.
- Eychene, A., Barnier, J.V., Apiou, F., Dutrillaux, B., and Calothy, G. 1992. Chromosomal assignment of two human B-raf(Rmil) proto-oncogene loci: B-raf-1 encoding the p94Braf/Rmil and B-raf-2, a processed pseudogene. *Oncogene* 7(8): 1657-1660.
- Finn, S.P., Smyth, P., Cahill, S., Streck, C., O'Regan, E.M., Flavin, R., Sherlock, J., Howells, D., Henfrey, R., Cullen, M. et al. 2007. Expression microarray analysis of papillary thyroid carcinoma and benign thyroid tissue: emphasis on the follicular variant and potential markers of malignancy. *Virchows Arch* 450(3): 249-260.
- Fisher, J.L., Mackie, P.S., Howard, M.L., Zhou, H., and Choong, P.F. 2001. The expression of the urokinase plasminogen activator system in metastatic murine osteosarcoma: an in vivo mouse model. *Clin Cancer Res* 7(6): 1654-1660.

Frame, M.C., Fincham, V.J., Carragher, N.O., and Wyke, J.A. 2002. v-Src's hold over actin and cell adhesions. *Nat Rev Mol Cell Biol* 3(4): 233-245.

Fransen, K., Klintenas, M., Osterstrom, A., Dimberg, J., Monstein, H.J., and Soderkvist, P. 2004. Mutation analysis of the BRAF, ARAF and RAF-1 genes in human colorectal adenocarcinomas. *Carcinogenesis* 25(4): 527-533.

Fukunaga-Kalabis, M., Martinez, G., Nguyen, T.K., Kim, D., Santiago-Walker, A., Roesch, A., and Herlyn, M. Tenascin-C promotes melanoma progression by maintaining the ABCB5-positive side population. *Oncogene*.

Furger, K.A., Allan, A.L., Wilson, S.M., Hota, C., Vantyghem, S.A., Postenka, C.O., Al-Katib, W., Chambers, A.F., and Tuck, A.B. 2003. Beta(3) integrin expression increases breast carcinoma cell responsiveness to the malignancy-enhancing effects of osteopontin. *Mol Cancer Res* 1(11): 810-819.

Galbraith, C.G., Yamada, K.M., and Sheetz, M.P. 2002. The relationship between force and focal complex development. *J Cell Biol* 159(4): 695-705.

Garnett, M.J. and Marais, R. 2004. Guilty as charged: B-RAF is a human oncogene. *Cancer Cell* 6(4): 313-319.

Garnett, M.J., Rana, S., Paterson, H., Barford, D., and Marais, R. 2005. Wild-type and mutant B-RAF activate C-RAF through distinct mechanisms involving heterodimerization. *Mol Cell* 20(6): 963-969.

Giehl, K. 2005. Oncogenic Ras in tumour progression and metastasis. *Biol Chem* 386(3): 193-205.

Gingras, D. and Beliveau, R. Emerging concepts in the regulation of membrane-type 1 matrix metalloproteinase activity. *Biochim Biophys Acta* 1803(1): 142-150.

Ginsberg, M.H., Partridge, A., and Shattil, S.J. 2005. Integrin regulation. *Curr Opin Cell Biol* 17(5): 509-516.

Giordano, T.J., Kuick, R., Thomas, D.G., Misek, D.E., Vinco, M., Sanders, D., Zhu, Z., Ciampi, R., Roh, M., Shedden, K. et al. 2005. Molecular classification of papillary thyroid carcinoma: distinct BRAF, RAS, and RET/PTC mutation-specific gene expression profiles discovered by DNA microarray analysis. *Oncogene* 24(44): 6646-6656.

Gire, V., Marshall, C.J., and Wynford-Thomas, D. 1999. Activation of mitogen-activated protein kinase is necessary but not sufficient for proliferation of human thyroid epithelial cells induced by mutant Ras. *Oncogene* 18(34): 4819-4832.

Goebeler, M., Kaufmann, D., Bocker, E.B., and Klein, C.E. 1996. Migration of highly aggressive melanoma cells on hyaluronic acid is associated with functional changes, increased turnover and shedding of CD44 receptors. *J Cell Sci* 109 ( Pt 7): 1957-1964.

Goley, E.D. and Welch, M.D. 2006. The ARP2/3 complex: an actin nucleator comes of age. *Nat Rev Mol Cell Biol* 7(10): 713-726.

Goodison, S., Urquidi, V., and Tarin, D. 1999. CD44 cell adhesion molecules. *Mol Pathol* 52(4): 189-196.

Green, J. A. and K. M. Yamada. 2007. Three-dimensional microenvironments modulate fibroblast signaling responses. *Adv Drug Deliv Rev* 59(13): 1293-1298.

Grieco, M., Cerrato, A., Santoro, M., Fusco, A., Melillo, R.M., and Vecchio, G. 1994. Cloning and characterization of H4 (D10S170), a gene involved in RET rearrangements in vivo. *Oncogene* 9(9): 2531-2535.

Guaita, S., Puig, I., Franci, C., Garrido, M., Dominguez, D., Batlle, E., Sancho, E., Dedhar, S., De Herreros, A.G., and Baulida, J. 2002. Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression. *J Biol Chem* 277(42): 39209-39216.

Guan, K.L., Figueroa, C., Brtva, T.R., Zhu, T., Taylor, J., Barber, T.D., and Vojtek, A.B. 2000. Negative regulation of the serine/threonine kinase B-Raf by Akt. *J Biol Chem* 275(35): 27354-27359.

Guarino, V., Faviana, P., Salvatore, G., Castellone, M.D., Cirafici, A.M., De Falco, V., Celetti, A., Giannini, R., Basolo, F., Melillo, R.M. et al. 2005. Osteopontin is overexpressed in human papillary thyroid carcinomas and enhances thyroid carcinoma cell invasiveness. *J Clin Endocrinol Metab* 90(9): 5270-5278.

Guo, Y.J., Liu, G., Wang, X., Jin, D., Wu, M., Ma, J., and Sy, M.S. 1994. Potential use of soluble CD44 in serum as indicator of tumor burden and metastasis in patients with gastric or colon cancer. *Cancer Res* 54(2): 422-426.

Hajra, K.M., Chen, D.Y., and Fearon, E.R. 2002. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res* 62(6): 1613-1618.

Hancock, J.F. 2003. Ras proteins: different signals from different locations. *Nat Rev Mol Cell Biol* 4(5): 373-384.

Hartsock, A. and Nelson, W.J. 2008. Adherens and tight junctions: structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta* 1778(3): 660-669.

Hardy, R.G., Vicente-Duenas, C., Gonzalez-Herrero, I., Anderson, C., Flores, T., Hughes, S., Tselepis, C., Ross, J.A., and Sanchez-Garcia, I. 2007. Snail family transcription factors are implicated in thyroid carcinogenesis. *Am J Pathol* 171(3): 1037-1046.

Hatsell, S., Rowlands, T., Hiremath, M., and Cowin, P. 2003. Beta-catenin and Tcfs in mammary development and cancer. *J Mammary Gland Biol Neoplasia* 8(2): 145-158.



Hickey, F.B., England, K., and Cotter, T.G. 2005. Bcr-Abl regulates osteopontin transcription via Ras, PI-3K, aPKC, Raf-1, and MEK. *J Leukoc Biol* 78(1): 289-300.

Hildebrand, J.D., Schaller, M.D., and Parsons, J.T. 1995. Paxillin, a tyrosine phosphorylated focal adhesion-associated protein binds to the carboxyl terminal domain of focal adhesion kinase. *Mol Biol Cell* 6(6): 637-647.

Hirata, H., Tatsumi, H., and Sokabe, M. 2008. Mechanical forces facilitate actin polymerization at focal adhesions in a zyxin-dependent manner. *J Cell Sci* 121(Pt 17): 2795-2804.

Hmitou, I., Druillennec, S., Valluet, A., Peyssonnaud, C., and Eychene, A. 2007. Differential regulation of B-raf isoforms by phosphorylation and autoinhibitory mechanisms. *Mol Cell Biol* 27(1): 31-43.

Hofmann, A., Laue, S., Rost, A.K., Scherbaum, W.A., and Aust, G. 1998. mRNA levels of membrane-type 1 matrix metalloproteinase (MT1-MMP), MMP-2, and MMP-9 and of their inhibitors TIMP-2 and TIMP-3 in normal thyrocytes and thyroid carcinoma cell lines. *Thyroid* 8(3): 203-214.

Huang, Y., Prasad, M., Lemon, W.J., Hampel, H., Wright, F.A., Kornacker, K., LiVolsi, V., Frankel, W., Kloos, R.T., Eng, C. et al. 2001. Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. *Proc Natl Acad Sci U S A* 98(26): 15044-15049.

Hugo, H., Ackland, M.L., Blick, T., Lawrence, M.G., Clements, J.A., Williams, E.D., and Thompson, E.W. 2007. Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. *J Cell Physiol* 213(2): 374-383.

Humphries, J.D., Wang, P., Streuli, C., Geiger, B., Humphries, M.J., and Ballestrem, C. 2007. Vinculin controls focal adhesion formation by direct interactions with talin and actin. *J Cell Biol* 179(5): 1043-1057.

Hunter, J.J., Tanaka, N., Rockman, H.A., Ross, J., Jr., and Chien, K.R. 1995. Ventricular expression of a MLC-2v-ras fusion gene induces cardiac hypertrophy and selective diastolic dysfunction in transgenic mice. *J Biol Chem* 270(39): 23173-23178.

Ichihara, M., Murakumo, Y., and Takahashi, M. 2004. RET and neuroendocrine tumors. *Cancer Lett* 204(2): 197-211.

Ikenouchi, J., Matsuda, M., Furuse, M., and Tsukita, S. 2003. Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. *J Cell Sci* 116(Pt 10): 1959-1967.

Ilmonen, S., Jahkola, T., Turunen, J.P., Muhonen, T., and Asko-Seljavaara, S. 2004. Tenascin-C in primary malignant melanoma of the skin. *Histopathology* 45(4): 405-411.

- Insall, R.H. and Machesky, L.M. 2009. Actin dynamics at the leading edge: from simple machinery to complex networks. *Dev Cell* 17(3): 310-322.
- Itoh, Y. and Seiki, M. 2004. MT1-MMP: an enzyme with multidimensional regulation. *Trends Biochem Sci* 29(6): 285-289.
- Ivanov, S.V., Ivanova, A.V., Goparaju, C.M., Chen, Y., Beck, A., and Pass, H.I. 2009. Tumorigenic properties of alternative osteopontin isoforms in mesothelioma. *Biochem Biophys Res Commun* 382(3): 514-518.
- Iwano, M., Plieth, D., Danoff, T.M., Xue, C., Okada, H., and Neilson, E.G. 2002. Evidence that fibroblasts derive from epithelium during tissue fibrosis. *J Clin Invest* 110(3): 341-350.
- Iyer, V., Pumiglia, K., and DiPersio, C.M. 2005. Alpha3beta1 integrin regulates MMP-9 mRNA stability in immortalized keratinocytes: a novel mechanism of integrin-mediated MMP gene expression. *J Cell Sci* 118(Pt 6): 1185-1195.
- Jain, S., Chakraborty, G., and Kundu, G.C. 2006. The crucial role of cyclooxygenase-2 in osteopontin-induced protein kinase C alpha/c-Src/IkappaB kinase alpha/beta-dependent prostate tumor progression and angiogenesis. *Cancer Res* 66(13): 6638-6648.
- Jamora, C., Lee, P., Kocieniewski, P., Azhar, M., Hosokawa, R., Chai, Y., and Fuchs, E. 2005. A signaling pathway involving TGF-beta2 and snail in hair follicle morphogenesis. *PLoS Biol* 3(1): e11.
- Janke, J., Schluter, K., Jandrig, B., Theile, M., Kolble, K., Arnold, W., Grinstein, E., Schwartz, A., Estevez-Schwarz, L., Schlag, P.M. et al. 2000. Suppression of tumorigenicity in breast cancer cells by the microfilament protein profilin 1. *J Exp Med* 191(10): 1675-1686.
- Jiang, S.M. 2000. The RET proto-oncogene in human cancers. *Oncogene* 19(49): 5590-5597.
- Jones, C.J., Kipling, D., Morris, M., Hepburn, P., Skinner, J., Bounacer, A., Wyllie, F.S., Ivan, M., Bartek, J., Wynford-Thomas, D. et al. 2000. Evidence for a telomere-independent "clock" limiting RAS oncogene-driven proliferation of human thyroid epithelial cells. *Mol Cell Biol* 20(15): 5690-5699.
- Jothy, S. 2003. CD44 and its partners in metastasis. *Clin Exp Metastasis* 20(3): 195-201.
- Kainz, C., Tempfer, C., Winkler, S., Sliutz, G., Koelbl, H., and Reinthaller, A. 1995. Serum CD44 splice variants in cervical cancer patients. *Cancer Lett* 90(2): 231-234.

Katagiri, Y.U., Sleeman, J., Fujii, H., Herrlich, P., Hotta, H., Tanaka, K., Chikuma, S., Yagita, H., Okumura, K., Murakami, M. et al. 1999. CD44 variants but not CD44s cooperate with beta1-containing integrins to permit cells to bind to osteopontin independently of arginine-glycine-aspartic acid, thereby stimulating cell motility and chemotaxis. *Cancer Res* 59(1): 219-226.

Kawano, Y., Okamoto, I., Murakami, D., Itoh, H., Yoshida, M., Ueda, S., and Saya, H. 2000. Ras oncoprotein induces CD44 cleavage through phosphoinositide 3-OH kinase and the rho family of small G proteins. *J Biol Chem* 275(38): 29628-29635.

Kebebew, E., Weng, J., Bauer, J., Ranvier, G., Clark, O.H., Duh, Q.Y., Shibru, D., Bastian, B., and Griffin, A. 2007. The prevalence and prognostic value of BRAF mutation in thyroid cancer. *Ann Surg* 246(3): 466-470; discussion 470-461.

Kimura, Y., Shiozaki, H., Hirao, M., Maeno, Y., Doki, Y., Inoue, M., Monden, T., Ando-Akatsuka, Y., Furuse, M., Tsukita, S. et al. 1997. Expression of occludin, tight-junction-associated protein, in human digestive tract. *Am J Pathol* 151(1): 45-54.

Kinsella, T.M. and Nolan, G.P. 1996. Episomal vectors rapidly and stably produce high-titer recombinant retrovirus. *Hum Gene Ther* 7(12): 1405-1413.

Kiosses, W.B., Shattil, S.J., Pampori, N., and Schwartz, M.A. 2001. Rac recruits high-affinity integrin alphavbeta3 to lamellipodia in endothelial cell migration. *Nat Cell Biol* 3(3): 316-320.

Klugbauer, S., Demidchik, E.P., Lengfelder, E., and Rabes, H.M. 1998. Detection of a novel type of RET rearrangement (PTC5) in thyroid carcinomas after Chernobyl and analysis of the involved RET-fused gene RFG5. *Cancer Res* 58(2): 198-203.

Klugbauer, S., Jauch, A., Lengfelder, E., Demidchik, E., and Rabes, H.M. 2000. A novel type of RET rearrangement (PTC8) in childhood papillary thyroid carcinomas and characterization of the involved gene (RFG8). *Cancer Res* 60(24): 7028-7032.

Klugbauer, S., Lengfelder, E., Demidchik, E.P., and Rabes, H.M. 1996. A new form of RET rearrangement in thyroid carcinomas of children after the Chernobyl reactor accident. *Oncogene* 13(5): 1099-1102.

Kominsky, S.L., Argani, P., Korz, D., Evron, E., Raman, V., Garrett, E., Rein, A., Sauter, G., Kallioniemi, O.P., and Sukumar, S. 2003. Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in situ and invasive ductal carcinoma of the breast. *Oncogene* 22(13): 2021-2033.

Kondo, T., Ezzat, S., and Asa, S.L. 2006. Pathogenetic mechanisms in thyroid follicular-cell neoplasia. *Nat Rev Cancer* 6(4): 292-306.

Kuphal, S. and Bosserhoff, A.K. 2006. Influence of the cytoplasmic domain of E-cadherin on endogenous N-cadherin expression in malignant melanoma. *Oncogene* 25(2): 248-259.

Kurrey, N.K., K, A., and Bapat, S.A. 2005. Snail and Slug are major determinants of ovarian cancer invasiveness at the transcription level. *Gynecol Oncol* 97(1): 155-165.

Lambertini, E., Lisignoli, G., Torreggiani, E., Manferdini, C., Gabusi, E., Franceschetti, T., Penolazzi, L., Gambari, R., Facchini, A., and Piva, R. 2009. Slug gene expression supports human osteoblast maturation. *Cell Mol Life Sci* 66(22): 3641-3653.

Le Bitoux, M.A. and Stamenkovic, I. 2008. Tumor-host interactions: the role of inflammation. *Histochem Cell Biol* 130(6): 1079-1090.

Le Clainche, C. and Carlier, M.F. 2008. Regulation of actin assembly associated with protrusion and adhesion in cell migration. *Physiol Rev* 88(2): 489-513.

Lee, J., Ishihara, A., Oxford, G., Johnson, B., and Jacobson, K. 1999. Regulation of cell movement is mediated by stretch-activated calcium channels. *Nature* 400(6742): 382-386.

Lee, J.L., Wang, M.J., Sudhir, P.R., Chen, G.D., Chi, C.W., and Chen, J.Y. 2007. Osteopontin promotes integrin activation through outside-in and inside-out mechanisms: OPN-CD44V interaction enhances survival in gastrointestinal cancer cells. *Cancer Res* 67(5): 2089-2097.

Lee, J.W., Soung, Y.H., Kim, S.Y., Park, W.S., Nam, S.W., Min, W.S., Kim, S.H., Lee, J.Y., Yoo, N.J., and Lee, S.H. 2005. Mutational analysis of the ARAF gene in human cancers. *APMIS* 113(1): 54-57.

Leroy, P. and Mostov, K.E. 2007. Slug is required for cell survival during partial epithelial-mesenchymal transition of HGF-induced tubulogenesis. *Mol Biol Cell* 18(5): 1943-1952.

Lewinski, A., Ferenc, T., Sporny, S., and Jarzab, B. 2000. Thyroid carcinoma: diagnostic and therapeutic approach; genetic background (review). *Endocr Regul* 34(2): 99-113.

Li, Y., Grenklo, S., Higgins, T., and Karlsson, R. 2008. The profilin:actin complex localizes to sites of dynamic actin polymerization at the leading edge of migrating cells and pathogen-induced actin tails. *Eur J Cell Biol* 87(11): 893-904.

Lin, J.D. and Chao, T.C. 2005. Vascular endothelial growth factor in thyroid cancers. *Cancer Biother Radiopharm* 20(6): 648-661.

Lin, Y.H., Huang, C.J., Chao, J.R., Chen, S.T., Lee, S.F., Yen, J.J., and Yang-Yen, H.F. 2000. Coupling of osteopontin and its cell surface receptor CD44 to the cell survival response elicited by interleukin-3 or granulocyte-macrophage colony-stimulating factor. *Mol Cell Biol* 20(8): 2734-2742.

Liu, D., Liu, Z., Condouris, S., and Xing, M. 2007. BRAF V600E maintains proliferation, transformation, and tumorigenicity of BRAF-mutant papillary thyroid cancer cells. *J Clin Endocrinol Metab* 92(6): 2264-2271.

Lorenz, M., Yamaguchi, H., Wang, Y., Singer, R.H., and Condeelis, J. 2004. Imaging sites of N-wasp activity in lamellipodia and invadopodia of carcinoma cells. *Curr Biol* 14(8): 697-703.

Lorenz, S., Vakonakis, I., Lowe, E.D., Campbell, I.D., Noble, M.E., and Hoellerer, M.K. 2008. Structural analysis of the interactions between paxillin LD motifs and alpha-parvin. *Structure* 16(10): 1521-1531.

Lubitz, C.C., Gallagher, L.A., Finley, D.J., Zhu, B., and Fahey, T.J., 3rd. 2005. Molecular analysis of minimally invasive follicular carcinomas by gene profiling. *Surgery* 138(6): 1042-1048; discussion 1048-1049.

Lubitz, C.C., Ugras, S.K., Kazam, J.J., Zhu, B., Scognamiglio, T., Chen, Y.T., and Fahey, T.J., 3rd. 2006. Microarray analysis of thyroid nodule fine-needle aspirates accurately classifies benign and malignant lesions. *J Mol Diagn* 8(4): 490-498; quiz 528.

Macpherson, I.R., Hooper, S., Serrels, A., McGarry, L., Ozanne, B.W., Harrington, K., Frame, M.C., Sahai, E., and Brunton, V.G. 2007. p120-catenin is required for the collective invasion of squamous cell carcinoma cells via a phosphorylation-independent mechanism. *Oncogene* 26(36): 5214-5228.

Maekawa, M., Ishizaki, T., Boku, S., Watanabe, N., Fujita, A., Iwamatsu, A., Obinata, T., Ohashi, K., Mizuno, K., and Narumiya, S. 1999. Signaling from Rho to the actin cytoskeleton through protein kinases ROCK and LIM-kinase. *Science* 285(5429): 895-898.

Marhaba, R. and Zoller, M. 2004. CD44 in cancer progression: adhesion, migration and growth regulation. *J Mol Histol* 35(3): 211-231.

Mason, C.S., Springer, C.J., Cooper, R.G., Superti-Furga, G., Marshall, C.J., and Marais, R. 1999. Serine and tyrosine phosphorylations cooperate in Raf-1, but not B-Raf activation. *EMBO J* 18(8): 2137-2148.

Mazzaferri, E.L. 1999. An overview of the management of papillary and follicular thyroid carcinoma. *Thyroid* 9(5): 421-427.

Medici, D., Hay, E.D., and Olsen, B.R. 2008. Snail and Slug promote epithelial-mesenchymal transition through beta-catenin-T-cell factor-4-dependent expression of transforming growth factor-beta3. *Mol Biol Cell* 19(11): 4875-4887.

Medico, E., Gentile, A., Lo Celso, C., Williams, T.A., Gambarotta, G., Trusolino, L., and Comoglio, P.M. 2001. Osteopontin is an autocrine mediator of hepatocyte growth factor-induced invasive growth. *Cancer Res* 61(15): 5861-5868.

Melillo, R.M., Castellone, M.D., Guarino, V., De Falco, V., Cirafici, A.M., Salvatore, G., Caiazzo, F., Basolo, F., Giannini, R., Kruhoffer, M. et al. 2005. The RET/PTC-RAS-BRAF linear signaling cascade mediates the motile and mitogenic phenotype of thyroid cancer cells. *J Clin Invest* 115(4): 1068-1081.

Mercer, K.E. and Pritchard, C.A. 2003. Raf proteins and cancer: B-Raf is identified as a mutational target. *Biochim Biophys Acta* 1653(1): 25-40.

Mete, O., Rotstein, L., and Asa, S.L. 2010. Controversies in thyroid pathology: thyroid capsule invasion and extrathyroidal extension. *Ann Surg Oncol* 17(2): 386-391.

Mi, Z., Guo, H., Wai, P.Y., Gao, C., and Kuo, P.C. 2006. Integrin-linked kinase regulates osteopontin-dependent MMP-2 and uPA expression to convey metastatic function in murine mammary epithelial cancer cells. *Carcinogenesis* 27(6): 1134-1145.

Mitsutake, N., Knauf, J.A., Mitsutake, S., Mesa, C., Jr., Zhang, L., and Fagin, J.A. 2005. Conditional BRAFV600E expression induces DNA synthesis, apoptosis, dedifferentiation, and chromosomal instability in thyroid PCCL3 cells. *Cancer Res* 65(6): 2465-2473.

Morgenstern, J.P. and Land, H. 1990. Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res* 18(12): 3587-3596.

Montero-Conde, C., Martin-Campos, J.M., Lerma, E., Gimenez, G., Martinez-Guitarte, J.L., Combalia, N., Montaner, D., Matias-Guiu, X., Dopazo, J., de Leiva, A. et al. 2008. Molecular profiling related to poor prognosis in thyroid carcinoma. Combining gene expression data and biological information. *Oncogene* 27(11): 1554-1561.

Mori, H., Tomari, T., Koshikawa, N., Kajita, M., Itoh, Y., Sato, H., Tojo, H., Yana, I., and Seiki, M. 2002. CD44 directs membrane-type 1 matrix metalloproteinase to lamellipodia by associating with its hemopexin-like domain. *EMBO J* 21(15): 3949-3959.

Mosimann, C., Hausmann, G., and Basler, K. 2009. Beta-catenin hits chromatin: regulation of Wnt target gene activation. *Nat Rev Mol Cell Biol* 10(4): 276-286.

Mullins, R.D., Heuser, J.A., and Pollard, T.D. 1998. The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proc Natl Acad Sci U S A* 95(11): 6181-6186.

Nagano, O. and Saya, H. 2004. Mechanism and biological significance of CD44 cleavage. *Cancer Sci* 95(12): 930-935.

Nagase, H., Visse, R., and Murphy, G. 2006. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc Res* 69(3): 562-573.

Nakahara, H., Howard, L., Thompson, E.W., Sato, H., Seiki, M., Yeh, Y., and Chen, W.T. 1997. Transmembrane/cytoplasmic domain-mediated membrane type 1-matrix metalloprotease docking to invadopodia is required for cell invasion. *Proc Natl Acad Sci U S A* 94(15): 7959-7964.

Nakamura, E.S., Koizumi, K., Kobayashi, M., and Saiki, I. 2004. Inhibition of lymphangiogenesis-related properties of murine lymphatic endothelial cells and lymph node metastasis of lung cancer by the matrix metalloproteinase inhibitor MMI270. *Cancer Sci* 95(1): 25-31.

Nakamura, H., Suenaga, N., Taniwaki, K., Matsuki, H., Yonezawa, K., Fujii, M., Okada, Y., and Seiki, M. 2004. Constitutive and induced CD44 shedding by ADAM-like proteases and membrane-type 1 matrix metalloproteinase. *Cancer Res* 64(3): 876-882.

Nakata, T., Kitamura, Y., Shimizu, K., Tanaka, S., Fujimori, M., Yokoyama, S., Ito, K., and Emi, M. 1999. Fusion of a novel gene, ELKS, to RET due to translocation t(10;12)(q11;p13) in a papillary thyroid carcinoma. *Genes Chromosomes Cancer* 25(2): 97-103.

Nakayama, H., Scott, I.C., and Cross, J.C. 1998. The transition to endoreduplication in trophoblast giant cells is regulated by the mSNA zinc finger transcription factor. *Dev Biol* 199(1): 150-163.

Namba, H., Nakashima, M., Hayashi, T., Hayashida, N., Maeda, S., Rogounovitch, T.I., Ohtsuru, A., Saenko, V.A., Kanematsu, T., and Yamashita, S. 2003. Clinical implication of hot spot BRAF mutation, V599E, in papillary thyroid cancers. *J Clin Endocrinol Metab* 88(9): 4393-4397.

Naor, D., Wallach-Dayana, S.B., Zahalka, M.A., and Sionov, R.V. 2008. Involvement of CD44, a molecule with a thousand faces, in cancer dissemination. *Semin Cancer Biol* 18(4): 260-267.

Nieto, M.A. 2002. The snail superfamily of zinc-finger transcription factors. *Nat Rev Mol Cell Biol* 3(3): 155-166.

Noren, N.K., Liu, B.P., Burridge, K., and Kreft, B. 2000. p120 catenin regulates the actin cytoskeleton via Rho family GTPases. *J Cell Biol* 150(3): 567-580.

Nunez, C. and Mendelsohn, G. 1989. Fine-needle aspiration and needle biopsy of the thyroid gland. *Pathol Annu* 24 Pt 1: 161-198.

Ochs, H.D. and Thrasher, A.J. 2006. The Wiskott-Aldrich syndrome. *J Allergy Clin Immunol* 117(4): 725-738; quiz 739.

Ohkubo, T. and Ozawa, M. 2004. The transcription factor Snail downregulates the tight junction components independently of E-cadherin downregulation. *J Cell Sci* 117(Pt 9): 1675-1685.

Okamoto, I., Kawano, Y., Matsumoto, M., Suga, M., Kaibuchi, K., Ando, M., and Saya, H. 1999a. Regulated CD44 cleavage under the control of protein kinase C, calcium influx, and the Rho family of small G proteins. *J Biol Chem* 274(36): 25525-25534.

Okamoto, I., Kawano, Y., Tsuiki, H., Sasaki, J., Nakao, M., Matsumoto, M., Suga, M., Ando, M., Nakajima, M., and Saya, H. 1999b. CD44 cleavage induced by a membrane-associated metalloprotease plays a critical role in tumor cell migration. *Oncogene* 18(7): 1435-1446.

Okamoto, I., Tsuiki, H., Kenyon, L.C., Godwin, A.K., Emlet, D.R., Holgado-Madruga, M., Lanham, I.S., Joynes, C.J., Vo, K.T., Guha, A. et al. 2002. Proteolytic cleavage of the CD44 adhesion molecule in multiple human tumors. *Am J Pathol* 160(2): 441-447.

Olmeda, D., Jorda, M., Peinado, H., Fabra, A., and Cano, A. 2007. Snail silencing effectively suppresses tumour growth and invasiveness. *Oncogene* 26(13): 1862-1874.

Omara-Opyene, A.L., Qiu, J., Shah, G.V., and Iczkowski, K.A. 2004. Prostate cancer invasion is influenced more by expression of a CD44 isoform including variant 9 than by Muc18. *Lab Invest* 84(7): 894-907.

O'Regan, A. 2003. The role of osteopontin in lung disease. *Cytokine Growth Factor Rev* 14(6): 479-488.

Palona, I., Namba, H., Mitsutake, N., Starenki, D., Podtcheko, A., Sedliarou, I., Ohtsuru, A., Saenko, V., Nagayama, Y., Umezawa, K. et al. 2006. BRAFV600E promotes invasiveness of thyroid cancer cells through nuclear factor kappaB activation. *Endocrinology* 147(12): 5699-5707.

Palyi-Krekk, Z., Barok, M., Kovacs, T., Saya, H., Nagano, O., Szollosi, J., and Nagy, P. 2008. EGFR and ErbB2 are functionally coupled to CD44 and regulate shedding, internalization and motogenic effect of CD44. *Cancer Lett* 263(2): 231-242.

Pankov, R., Cukierman, E., Katz, B.Z., Matsumoto, K., Lin, D.C., Lin, S., Hahn, C., and Yamada, K.M. 2000. Integrin dynamics and matrix assembly: tensin-dependent translocation of alpha(5)beta(1) integrins promotes early fibronectin fibrillogenesis. *J Cell Biol* 148(5): 1075-1090.

Pankov, R., Y. Endo 2005. A Rac switch regulates random versus directionally persistent cell migration. *J Cell Biol* 170(5): 793-802.

Peinado, H., Olmeda, D., and Cano, A. 2007. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* 7(6): 415-428.

Peinado, H., Portillo, F., and Cano, A. 2004. Transcriptional regulation of cadherins during development and carcinogenesis. *Int J Dev Biol* 48(5-6): 365-375.



Peinado, H., Quintanilla, M., and Cano, A. 2003. Transforming growth factor beta-1 induces snail transcription factor in epithelial cell lines: mechanisms for epithelial mesenchymal transitions. *J Biol Chem* 278(23): 21113-21123.

Pfaff, M., Liu, S., Erle, D.J., and Ginsberg, M.H. 1998. Integrin beta cytoplasmic domains differentially bind to cytoskeletal proteins. *J Biol Chem* 273(11): 6104-6109.

Philip, S., Bulbule, A., and Kundu, G.C. 2001. Osteopontin stimulates tumor growth and activation of promatrix metalloproteinase-2 through nuclear factor-kappa B-mediated induction of membrane type 1 matrix metalloproteinase in murine melanoma cells. *J Biol Chem* 276(48): 44926-44935.

Ponta, H., Sherman, L., and Herrlich, P.A. 2003. CD44: from adhesion molecules to signalling regulators. *Nat Rev Mol Cell Biol* 4(1): 33-45.

Popoff, M.R. and Geny, B. 2009. Multifaceted role of Rho, Rac, Cdc42 and Ras in intercellular junctions, lessons from toxins. *Biochim Biophys Acta* 1788(4): 797-812.

Prasad, M.L., Pellegata, N.S., Huang, Y., Nagaraja, H.N., de la Chapelle, A., and Kloos, R.T. 2005. Galectin-3, fibronectin-1, CITED-1, HBME1 and cytokeratin-19 immunohistochemistry is useful for the differential diagnosis of thyroid tumors. *Mod Pathol* 18(1): 48-57.

Prensner, J.R. and Chinnaiyan, A.M. 2009. Oncogenic gene fusions in epithelial carcinomas. *Curr Opin Genet Dev* 19(1): 82-91.

Putzer, B.M. and Drosten, M. 2004. The RET proto-oncogene: a potential target for molecular cancer therapy. *Trends Mol Med* 10(7): 351-357.

Ramaiah, S.K. and Rittling, S. 2008. Pathophysiological role of osteopontin in hepatic inflammation, toxicity, and cancer. *Toxicol Sci* 103(1): 4-13.

Rangaswami, H., Bulbule, A., and Kundu, G.C. 2006. Osteopontin: role in cell signaling and cancer progression. *Trends Cell Biol* 16(2): 79-87.

Rapp, U.R., Goldsborough, M.D., Mark, G.E., Bonner, T.I., Groffen, J., Reynolds, F.H., Jr., and Stephenson, J.R. 1983. Structure and biological activity of v-raf, a unique oncogene transduced by a retrovirus. *Proc Natl Acad Sci U S A* 80(14): 4218-4222.

Renault, M.A., Jalvy, S., Belloc, I., Pasquet, S., Sena, S., Olive, M., Desgranges, C., and Gadeau, A.P. 2003. AP-1 is involved in UTP-induced osteopontin expression in arterial smooth muscle cells. *Circ Res* 93(7): 674-681.

Riento, K. and Ridley, A.J. 2003. Rocks: multifunctional kinases in cell behaviour. *Nat Rev Mol Cell Biol* 4(6): 446-456.

Ristamaki, R., Joensuu, H., Lappalainen, K., Teerenhovi, L., and Jalkanen, S. 1997. Elevated serum CD44 level is associated with unfavorable outcome in non-Hodgkin's lymphoma. *Blood* 90(10): 4039-4045.

Robertson, B.W. and Chellaiah, M.A. 2010. Osteopontin induces beta-catenin signaling through activation of Akt in prostate cancer cells. *Exp Cell Res* 316(1): 1-11.

Rohatgi, R., Ma, L., Miki, H., Lopez, M., Kirchhausen, T., Takenawa, T., and Kirschner, M.W. 1999. The interaction between N-WASP and the Arp2/3 complex links Cdc42-dependent signals to actin assembly. *Cell* 97(2): 221-231.

Rohde, F., Rimkus, C., Friederichs, J., Rosenberg, R., Marthen, C., Doll, D., Holzmann, B., Siewert, J.R., and Janssen, K.P. 2007. Expression of osteopontin, a target gene of de-regulated Wnt signaling, predicts survival in colon cancer. *Int J Cancer* 121(8): 1717-1723.

Roskelley, C. D. and M. J. Bissell. 1995. Dynamic reciprocity revisited: a continuous, bidirectional flow of information between cells and the extracellular matrix regulates mammary epithelial cell function. *Biochem Cell Biol* 73(7-8): 391-397.

Rupp, P.A., Visconti, R.P., Czirok, A., Cheresch, D.A., and Little, C.D. 2008. Matrix metalloproteinase 2-integrin alpha(v)beta3 binding is required for mesenchymal cell invasive activity but not epithelial locomotion: a computational time-lapse study. *Mol Biol Cell* 19(12): 5529-5540.

Saenko, V., Rogounovitch, T., Shimizu-Yoshida, Y., Abrosimov, A., Lushnikov, E., Roumiantsev, P., Matsumoto, N., Nakashima, M., Meirmanov, S., Ohtsuru, A. et al. 2003. Novel tumorigenic rearrangement, Delta rfp/ret, in a papillary thyroid carcinoma from externally irradiated patient. *Mutat Res* 527(1-2): 81-90.

Sakai, T., Li, S., Docheva, D., Grashoff, C., Sakai, K., Kostka, G., Braun, A., Pfeifer, A., Yurchenco, P.D., and Fassler, R. 2003. Integrin-linked kinase (ILK) is required for polarizing the epiblast, cell adhesion, and controlling actin accumulation. *Genes Dev* 17(7): 926-940.

Salassidis, K., Bruch, J., Zitzelsberger, H., Lengfelder, E., Kellerer, A.M., and Bauchinger, M. 2000. Translocation t(10;14)(q11.2;q22.1) fusing the kinetin to the RET gene creates a novel rearranged form (PTC8) of the RET proto-oncogene in radiation-induced childhood papillary thyroid carcinoma. *Cancer Res* 60(11): 2786-2789.

Samanna, V., Wei, H., Ego-Osuala, D., and Chellaiah, M.A. 2006. Alpha-V-dependent outside-in signaling is required for the regulation of CD44 surface expression, MMP-2 secretion, and cell migration by osteopontin in human melanoma cells. *Exp Cell Res* 312(12): 2214-2230.

Santoro, M., Dathan, N.A., Berlingieri, M.T., Bongarzone, I., Paulin, C., Grieco, M., Pierotti, M.A., Vecchio, G., and Fusco, A. 1994. Molecular characterization of RET/PTC3; a novel rearranged version of the RET proto-oncogene in a human thyroid papillary carcinoma. *Oncogene* 9(2): 509-516.

Sakakibara, A., Furuse, M., Saitou, M., Ando-Akatsuka, Y., and Tsukita, S. 1997. Possible involvement of phosphorylation of occludin in tight junction formation. *J Cell Biol* 137(6): 1393-1401.

Sato, H., Takino, T., Okada, Y., Cao, J., Shinagawa, A., Yamamoto, E., and Seiki, M. 1994. A matrix metalloproteinase expressed on the surface of invasive tumour cells. *Nature* 370(6484): 61-65.

Saunders, R.M., Holt, M.R., Jennings, L., Sutton, D.H., Barsukov, I.L., Bobkov, A., Liddington, R.C., Adamson, E.A., Dunn, G.A., and Critchley, D.R. 2006. Role of vinculin in regulating focal adhesion turnover. *Eur J Cell Biol* 85(6): 487-500.

Savagner, P., Yamada, K.M., and Thiery, J.P. 1997. The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. *J Cell Biol* 137(6): 1403-1419.

Schaller, M.D., Otey, C.A., Hildebrand, J.D., and Parsons, J.T. 1995. Focal adhesion kinase and paxillin bind to peptides mimicking beta integrin cytoplasmic domains. *J Cell Biol* 130(5): 1181-1187.

Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F., and Pachnis, V. 1994. Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. *Nature* 367(6461): 380-383.

Schwartz, M. 2004. Rho signalling at a glance. *J Cell Sci* 117(Pt 23): 5457-5458.

Sefton, M., Sanchez, S., and Nieto, M.A. 1998. Conserved and divergent roles for members of the Snail family of transcription factors in the chick and mouse embryo. *Development* 125(16): 3111-3121.

Selmecki, D., S. Mosler. 2005. Cell motility as persistent random motion: theories from experiments. *Biophys J* 89(2): 912-931.

Senger, D.R., Perruzzi, C.A., Gracey, C.F., Papadopoulos, A., and Tenen, D.G. 1988. Secreted phosphoproteins associated with neoplastic transformation: close homology with plasma proteins cleaved during blood coagulation. *Cancer Res* 48(20): 5770-5774.

Senger, D.R., Perruzzi, C.A., Papadopoulos, A., and Tenen, D.G. 1989. Purification of a human milk protein closely similar to tumor-secreted phosphoproteins and osteopontin. *Biochim Biophys Acta* 996(1-2): 43-48.

Senger, D.R., Wirth, D.F., and Hynes, R.O. 1979. Transformed mammalian cells secrete specific proteins and phosphoproteins. *Cell* 16(4): 885-893.

Shattil, S.J., Kim, C., and Ginsberg, M.H. 2010. The final steps of integrin activation: the end game. *Nat Rev Mol Cell Biol* 11(4): 288-300.

Sherman, S.I. 2003. Thyroid carcinoma. *Lancet* 361(9356): 501-511.

Shih, J.Y., Tsai, M.F., Chang, T.H., Chang, Y.L., Yuan, A., Yu, C.J., Lin, S.B., Liou, G.Y., Lee, M.L., Chen, J.J. et al. 2005. Transcription repressor slug promotes carcinoma invasion and predicts outcome of patients with lung adenocarcinoma. *Clin Cancer Res* 11(22): 8070-8078.

Shin, E., Hong, S.W., Kim, S.H., and Yang, W.I. 2004. Expression of down stream molecules of RET (p-ERK, p-p38 MAPK, p-JNK and p-AKT) in papillary thyroid carcinomas. *Yonsei Med J* 45(2): 306-313.

Shinohara, M.L., Kim, H.J., Kim, J.H., Garcia, V.A., and Cantor, H. 2008. Alternative translation of osteopontin generates intracellular and secreted isoforms that mediate distinct biological activities in dendritic cells. *Proc Natl Acad Sci U S A* 105(20): 7235-7239.

Sivertsen, S., Hadar, R., Elloul, S., Vintman, L., Bedrossian, C., Reich, R., and Davidson, B. 2006. Expression of Snail, Slug and Sip1 in malignant mesothelioma effusions is associated with matrix metalloproteinase, but not with cadherin expression. *Lung Cancer* 54(3): 309-317.

Smith, L.L., Greenfield, B.W., Aruffo, A., and Giachelli, C.M. 1999. CD44 is not an adhesive receptor for osteopontin. *J Cell Biochem* 73(1): 20-30.

Sneath, R.J. and Mangham, D.C. 1998. The normal structure and function of CD44 and its role in neoplasia. *Mol Pathol* 51(4): 191-200.

Soares, P., Trovisco, V., Rocha, A.S., Lima, J., Castro, P., Preto, A., Maximo, V., Botelho, T., Seruca, R., and Sobrinho-Simoes, M. 2003. BRAF mutations and RET/PTC rearrangements are alternative events in the etiopathogenesis of PTC. *Oncogene* 22(29): 4578-4580.

Sounni, N.E., Devy, L., Hajitou, A., Frankenne, F., Munaut, C., Gilles, C., Deroanne, C., Thompson, E.W., Foidart, J.M., and Noel, A. 2002. MT1-MMP expression promotes tumor growth and angiogenesis through an up-regulation of vascular endothelial growth factor expression. *FASEB J* 16(6): 555-564.

Sozzi, G., Bongarzone, I., Miozzo, M., Borrello, M.G., Blutti, M.G., Pilotti, S., Della Porta, G., and Pierotti, M.A. 1994. A t(10;17) translocation creates the RET/PTC2 chimeric transforming sequence in papillary thyroid carcinoma. *Genes Chromosomes Cancer* 9(4): 244-250.

Staibano, S., Merolla, F., Testa, D., Iovine, R., Mascolo, M., Guarino, V., Castellone, M.D., Di Benedetto, M., Galli, V., Motta, S. et al. 2007. OPN/CD44v6 overexpression in laryngeal dysplasia and correlation with clinical outcome. *Br J Cancer* 97(11): 1545-1551.

Steiner, H., Fluhrer, R., and Haass, C. 2008. Intramembrane proteolysis by gamma-secretase. *J Biol Chem* 283(44): 29627-29631.

Stolf, B.S., Santos, M.M., Simao, D.F., Diaz, J.P., Cristo, E.B., Hirata, R., Jr., Curado, M.P., Neves, E.J., Kowalski, L.P., and Carvalho, A.F. 2006. Class distinction between follicular adenomas and follicular carcinomas of the thyroid gland on the basis of their signature expression. *Cancer* 106(9): 1891-1900.

Strongin, A.Y., Collier, I., Bannikov, G., Marmer, B.L., Grant, G.A., and Goldberg, G.I. 1995. Mechanism of cell surface activation of 72-kDa type IV collagenase. Isolation of the activated form of the membrane metalloprotease. *J Biol Chem* 270(10): 5331-5338.

Subauste, M.C., Pertz, O., Adamson, E.D., Turner, C.E., Junger, S., and Hahn, K.M. 2004. Vinculin modulation of paxillin-FAK interactions regulates ERK to control survival and motility. *J Cell Biol* 165(3): 371-381.

Sugden, P.H. and Clerk, A. 2000. Activation of the small GTP-binding protein Ras in the heart by hypertrophic agonists. *Trends Cardiovasc Med* 10(1): 1-8.

Sugimachi, K., Tanaka, S., Kameyama, T., Taguchi, K., Aishima, S., Shimada, M., and Tsuneyoshi, M. 2003. Transcriptional repressor snail and progression of human hepatocellular carcinoma. *Clin Cancer Res* 9(7): 2657-2664.

Sumi, T., Matsumoto, K., Takai, Y., and Nakamura, T. 1999. Cofilin phosphorylation and actin cytoskeletal dynamics regulated by rho- and Cdc42-activated LIM-kinase 2. *J Cell Biol* 147(7): 1519-1532.

Surh, Y.J. and Kundu, J.K. 2005. Signal transduction network leading to COX-2 induction: a road map in search of cancer chemopreventives. *Arch Pharm Res* 28(1): 1-15.

Suster, S. 2006. Thyroid tumors with a follicular growth pattern: problems in differential diagnosis. *Arch Pathol Lab Med* 130(7): 984-988.

Svitkina, T.M. and Borisy, G.G. 1999. Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *J Cell Biol* 145(5): 1009-1026.

Tadokoro, S., Shattil, S.J., Eto, K., Tai, V., Liddington, R.C., de Pereda, J.M., Ginsberg, M.H., and Calderwood, D.A. 2003. Talin binding to integrin beta tails: a final common step in integrin activation. *Science* 302(5642): 103-106.

Takahashi, M. 1988. Structure and expression of the ret transforming gene. *IARC Sci Publ*(92): 189-197.

Takai, Y., Sasaki, T., and Matozaki, T. 2001. Small GTP-binding proteins. *Physiol Rev* 81(1): 153-208.

Teramoto, H., Castellone, M.D., Malek, R.L., Letwin, N., Frank, B., Gutkind, J.S., and Lee, N.H. 2005. Autocrine activation of an osteopontin-CD44-Rac pathway enhances invasion and transformation by H-RasV12. *Oncogene* 24(3): 489-501.

Tezuka, K., Denhardt, D.T., Rodan, G.A., and Harada, S. 1996. Stimulation of mouse osteopontin promoter by v-Src is mediated by a CCAAT box-binding factor. *J Biol Chem* 271(37): 22713-22717.

Tong, Q., Xing, S., and Jhiang, S.M. 1997. Leucine zipper-mediated dimerization is essential for the PTC1 oncogenic activity. *J Biol Chem* 272(14): 9043-9047.

Tripathi, M.K., Misra, S., and Chaudhuri, G. 2005. Negative regulation of the expressions of cytokeratins 8 and 19 by SLUG repressor protein in human breast cells. *Biochem Biophys Res Commun* 329(2): 508-515.

Trovisco, V., Vieira de Castro, I., Soares, P., Maximo, V., Silva, P., Magalhaes, J., Abrosimov, A., Guiu, X.M., and Sobrinho-Simoes, M. 2004. BRAF mutations are associated with some histological types of papillary thyroid carcinoma. *J Pathol* 202(2): 247-251.

Tseleni-Balafouta, S., Gakiopoulou, H., Fanourakis, G., Voutsinas, G., Balafoutas, D., and Patsouris, E. 2006. Tenascin-C protein expression and mRNA splice variants in thyroid carcinoma. *Exp Mol Pathol* 80(2): 177-182

Tuck, A.B., Hota, C., and Chambers, A.F. 2001. Osteopontin(OPN)-induced increase in human mammary epithelial cell invasiveness is urokinase (uPA)-dependent. *Breast Cancer Res Treat* 70(3): 197-204.

Turner, F.E., Broad, S., Khanim, F.L., Jeanes, A., Talma, S., Hughes, S., Tselepis, C., and Hotchin, N.A. 2006. Slug regulates integrin expression and cell proliferation in human epidermal keratinocytes. *J Biol Chem* 281(30): 21321-21331.

Uchikado, Y., Natsugoe, S., Okumura, H., Setoyama, T., Matsumoto, M., Ishigami, S., and Aikou, T. 2005. Slug Expression in the E-cadherin preserved tumors is related to prognosis in patients with esophageal squamous cell carcinoma. *Clin Cancer Res* 11(3): 1174-1180.

Ugolini, C., Giannini, R., Lupi, C., Salvatore, G., Miccoli, P., Proietti, A., Elisei, R., Santoro, M., and Basolo, F. 2007. Presence of BRAF V600E in very early stages of papillary thyroid carcinoma. *Thyroid* 17(5): 381-388.

Urano, T., Liu, J., Zhang, P., Fan, Y., Egile, C., Li, R., Mueller, S.C., and Zhan, X. 2001. Activation of Arp2/3 complex-mediated actin polymerization by cortactin. *Nat Cell Biol* 3(3): 259-266

van der Voort, R., Taher, T.E., Wielenga, V.J., Spaargaren, M., Prevo, R., Smit, L., David, G., Hartmann, G., Gherardi, E., and Pals, S.T. 1999. Heparan sulfate-modified CD44 promotes hepatocyte growth factor/scatter factor-induced signal transduction through the receptor tyrosine kinase c-Met. *J Biol Chem* 274(10): 6499-6506.

van Staveren, W.C., Solis, D.W., Delys, L., Duprez, L., Andry, G., Franc, B., Thomas, G., Libert, F., Dumont, J.E., Detours, V. et al. 2007. Human thyroid tumor

cell lines derived from different tumor types present a common dedifferentiated phenotype. *Cancer Res* 67(17): 8113-8120.

Vasko, V., Espinosa, A.V., Scouten, W., He, H., Auer, H., Liyanarachchi, S., Larin, A., Savchenko, V., Francis, G.L., de la Chapelle, A. et al. 2007. Gene expression and functional evidence of epithelial-to-mesenchymal transition in papillary thyroid carcinoma invasion. *Proc Natl Acad Sci U S A* 104(8): 2803-2808.

Vasko, V., Ferrand, M., Di Cristofaro, J., Carayon, P., Henry, J.F., and de Micco, C. 2003. Specific pattern of RAS oncogene mutations in follicular thyroid tumors. *J Clin Endocrinol Metab* 88(6): 2745-2752.

Vicovac, L. and Aplin, J.D. 1996. Epithelial-mesenchymal transition during trophoblast differentiation. *Acta Anat (Basel)* 156(3): 202-216.

Visse, R. and Nagase, H. 2003. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res* 92(8): 827-839.

Wai, P.Y. and Kuo, P.C. 2004. The role of Osteopontin in tumor metastasis. *J Surg Res* 121(2): 228-241.

Wai, P.Y. and Kuo, P.C. 2008. Osteopontin: regulation in tumor metastasis. *Cancer Metastasis Rev* 27(1): 103-118.

Wallace, J.M. 2002. Nutritional and botanical modulation of the inflammatory cascade--eicosanoids, cyclooxygenases, and lipoxygenases--as an adjunct in cancer therapy. *Integr Cancer Ther* 1(1): 7-37; discussion 37.

Wang, Z., Wade, P., Mandell, K.J., Akyildiz, A., Parkos, C.A., Mrsny, R.J., and Nusrat, A. 2007. Raf 1 represses expression of the tight junction protein occludin via activation of the zinc-finger transcription factor slug. *Oncogene* 26(8): 1222-1230.

Weber, C.K., Slupsky, J.R., Kalmes, H.A., and Rapp, U.R. 2001. Active Ras induces heterodimerization of cRaf and BRaf. *Cancer Res* 61(9): 3595-3598.

Weber, F., Shen, L., Aldred, M.A., Morrison, C.D., Frilling, A., Saji, M., Schuppert, F., Broelsch, C.E., Ringel, M.D., and Eng, C. 2005. Genetic classification of benign and malignant thyroid follicular neoplasia based on a three-gene combination. *J Clin Endocrinol Metab* 90(5): 2512-2521.

Welch, M.D. 1999. The world according to Arp: regulation of actin nucleation by the Arp2/3 complex. *Trends Cell Biol* 9(11): 423-427.

Welch, M.D. and Mullins, R.D. 2002. Cellular control of actin nucleation. *Annu Rev Cell Dev Biol* 18: 247-288.

Westermarck, K., Lundqvist, M., Wallin, G., Dahlman, T., Hacker, G.W., Heldin, N.E., and Grimelius, L. 1996. EGF-receptors in human normal and pathological thyroid tissue. *Histopathology* 28(3): 221-227.

Wickstrom, S.A., Lange, A., Montanez, E., and Fassler, R. The ILK/PINCH/parvin complex: the kinase is dead, long live the pseudokinase! *EMBO J* 29(2): 281-291.

Wielenga, V.J., Smits, R., Korinek, V., Smit, L., Kielman, M., Fodde, R., Clevers, H., and Pals, S.T. 1999. Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. *Am J Pathol* 154(2): 515-523

Wiesner, S., Legate, K.R., and Fassler, R. 2005. Integrin-actin interactions. *Cell Mol Life Sci* 62(10): 1081-1099.

Williams, D.W., Williams, E.D., and Wynford-Thomas, D. 1988. Loss of dependence on IGF-1 for proliferation of human thyroid adenoma cells. *Br J Cancer* 57(6): 535-539.

Woods, A.J., White, D.P., Caswell, P.T., and Norman, J.C. 2004. PKD1/PKCmu promotes alphavbeta3 integrin recycling and delivery to nascent focal adhesions. *EMBO J* 23(13): 2531-2543.

Wozniak, M.A., Modzelewska, K., Kwong, L., and Keely, P.J. 2004. Focal adhesion regulation of cell behavior. *Biochim Biophys Acta* 1692(2-3): 103-119.

Wu, W.S., Heinrichs, S., Xu, D., Garrison, S.P., Zambetti, G.P., Adams, J.M., and Look, A.T. 2005. Slug antagonizes p53-mediated apoptosis of hematopoietic progenitors by repressing puma. *Cell* 123(4): 641-653.

Wu, X., Suetsugu, S., Cooper, L.A., Takenawa, T., and Guan, J.L. 2004. Focal adhesion kinase regulation of N-WASP subcellular localization and function. *J Biol Chem* 279(10): 9565-9576.

Wymann, M.P., Zvelebil, M., and Laffargue, M. 2003. Phosphoinositide 3-kinase signalling--which way to target? *Trends Pharmacol Sci* 24(7): 366-376.

Yamaguchi, H. and Condeelis, J. 2007. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochim Biophys Acta* 1773(5): 642-652.

Yamazaki, D., Fujiwara, T., Suetsugu, S., and Takenawa, T. 2005. A novel function of WAVE in lamellipodia: WAVE1 is required for stabilization of lamellipodial protrusions during cell spreading. *Genes Cells* 10(5): 381-392.

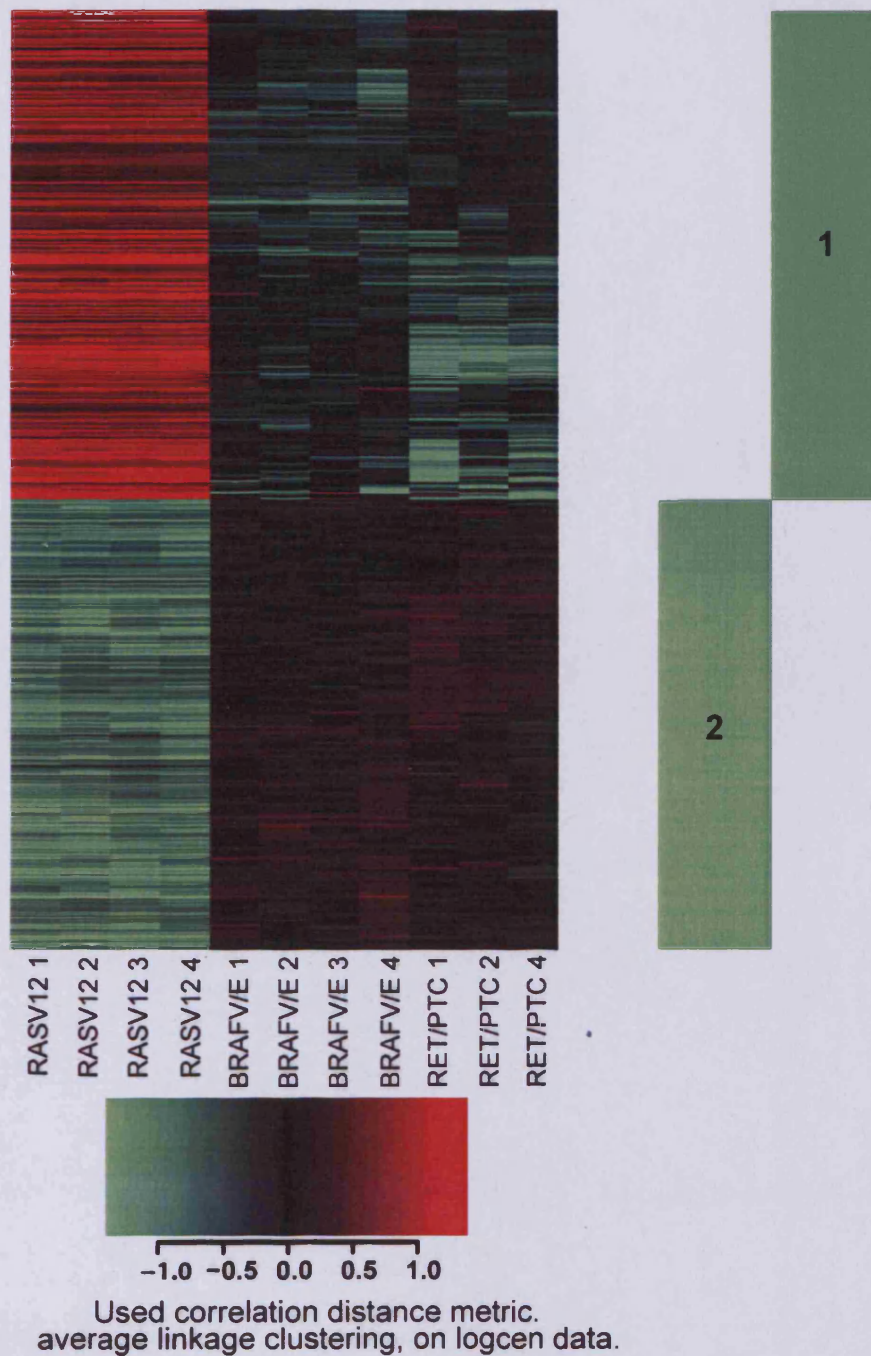
Yang, H.W., Menon, L.G., Black, P.M., Carroll, R.S., and Johnson, M.D. 2010. SNAI2/Slug promotes growth and invasion in human gliomas. *BMC Cancer* 10: 301.

Yokoyama, K., Kamata, N., Fujimoto, R., Tsutsumi, S., Tomonari, M., Taki, M., Hosokawa, H., and Nagayama, M. 2003. Increased invasion and matrix metalloproteinase-2 expression by Snail-induced mesenchymal transition in squamous cell carcinomas. *Int J Oncol* 22(4): 891-898.

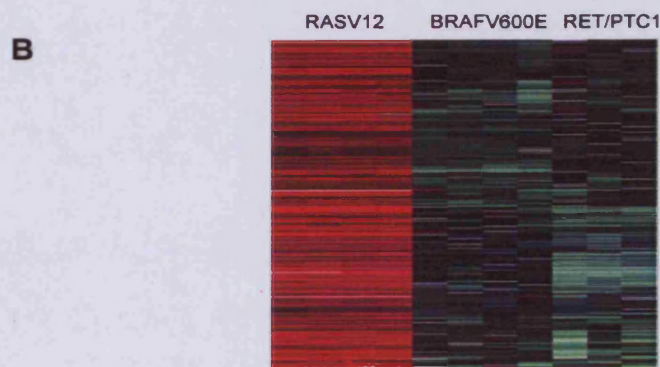


- Yoshimi, R., Yamaji, S., Suzuki, A., Mishima, W., Okamura, M., Obana, T., Matsuda, C., Miwa, Y., Ohno, S., and Ishigatsubo, Y. 2006. The gamma-parvin-integrin-linked kinase complex is critically involved in leukocyte-substrate interaction. *J Immunol* 176(6): 3611-3624.
- Yu, Q., Toole, B.P., and Stamenkovic, I. 1997. Induction of apoptosis of metastatic mammary carcinoma cells in vivo by disruption of tumor cell surface CD44 function. *J Exp Med* 186(12): 1985-1996.
- Zhao, J., Leonard, C., Brunner, E., Genssenjager, E., Heitz, P.U., and Odermatt, B. 2006. Molecular characterization of well-differentiated human thyroid carcinomas by cDNA arrays. *Int J Oncol* 29(5): 1041-1051.
- Zhao, J., Leonard, C., Genssenjager, E., Heitz, P.U., Moch, H., and Odermatt, B. 2008. Differentiation of human follicular thyroid adenomas from carcinomas by gene expression profiling. *Oncol Rep* 19(2): 329-337.
- Zhu, B., Suzuki, K., Goldberg, H.A., Rittling, S.R., Denhardt, D.T., McCulloch, C.A., and Sodek, J. 2004. Osteopontin modulates CD44-dependent chemotaxis of peritoneal macrophages through G-protein-coupled receptors: evidence of a role for an intracellular form of osteopontin. *J Cell Physiol* 198(1): 155-167.
- Zhu, Y., Denhardt, D.T., Cao, H., Sutphin, P.D., Koong, A.C., Giaccia, A.J., and Le, Q.T. 2005. Hypoxia upregulates osteopontin expression in NIH-3T3 cells via a Ras-activated enhancer. *Oncogene* 24(43): 6555-6563.
- Ziegler, W.H., Liddington, R.C., and Critchley, D.R. 2006. The structure and regulation of vinculin. *Trends Cell Biol* 16(9): 453-460.
- Zohar, R., Suzuki, N., Suzuki, K., Arora, P., Glogauer, M., McCulloch, C.A., and Sodek, J. 2000. Intracellular osteopontin is an integral component of the CD44-ERM complex involved in cell migration. *J Cell Physiol* 184(1): 118-130.
- Zou, M., Al-Baradie, R.S., Al-Hindi, H., Farid, N.R., and Shi, Y. 2005. S100A4 (Mts1) gene overexpression is associated with invasion and metastasis of papillary thyroid carcinoma. *Br J Cancer* 93(11): 1277-1284.
- Zou, L., Jaramillo, M., Whaley, D., Wells, A., Panchapakesa, V., Das, T., and Roy, P. 2007. Profilin-1 is a negative regulator of mammary carcinoma aggressiveness. *Br J Cancer* 97(10): 1361-1371.

## Appendix

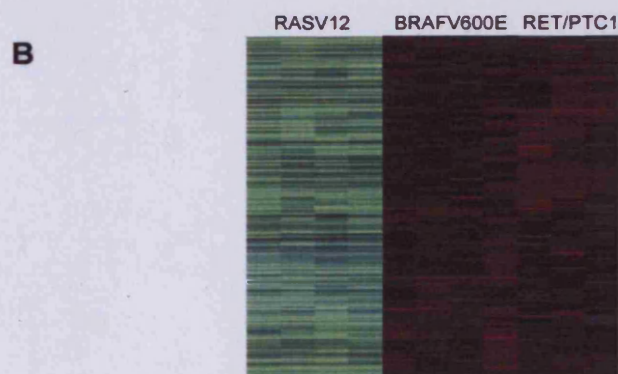


**Figure 1.** *t*-test analysis of genotype-specific changes in gene expression. Hierarchical clustering displaying the genotype specific transcriptional changes of 705 differentially regulated probe sets in RASV12 compared to BRAF<sup>V600E</sup> and RET/PTC1 oncogene infected thyrocytes. The columns correspond to the oncogene and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation.



**Figure 2.** Enrichment analysis of pathway folders identified in t-test cluster 1. The enrichment method was subjected to an FDR of 0.05. **A**, The order of map folders is based on their significance in the cluster. The dark orange bars dictate significance set at 0.05%. Although the light/orange bars are displayed they are not discussed due to the possibility of false positive results. For the t-test cluster 1, there were 33 map folders identified with 10 of significance. **B**, The corresponding heatmap of the t-test cluster 1, displaying an overall expression pattern of 334 up-regulated probes in RASV12 and median/down-regulation in both BRAF<sup>V600E</sup> and RET/PTC1. The columns correspond to the oncogenic condition and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation, whilst median expression is signified by black.





**Figure 3.** Enrichment analysis of pathway folders identified in t-test cluster 2. The enrichment method was subjected to an FDR of 0.05. **A**, The order of map folders is based on their significance in the cluster. The dark orange bars dictate significance set at 0.05%. Although the light/orange bars are displayed they are not discussed due to the possibility of false positive results. For the t-test cluster 2, there were 33 map folders identified with 5 of significance. **B**, The corresponding heatmap of the t-test cluster 2, displaying an overall expression pattern of 334 down-regulated probes in RASV12 and median/up-regulation in both BRAF<sup>V600E</sup> and RET/PTC1. The columns correspond to the oncogenic condition and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation, whilst median expression is signified by black.

**Table 1.0 T-test cluster 1**

#	Gene Symbol	Protein name	Object type
1	<u>ACP1</u>	Low molecular weight phosphotyrosine protein phosphatase	Protein phosphatase
2	<u>ACSL3</u>	Long-chain-fatty-acid--CoA ligase 3	Generic enzyme
3	<u>ADM</u>	ADM	Receptor ligand
4	<u>AE01</u>	Neuroblastoma breakpoint family member 9	Generic protein
5	<u>AK3L1</u>	Adenylate kinase isoenzyme 4, mitochondrial	Generic kinase
6	<u>AKAP10</u>	A kinase anchor protein 10, mitochondrial	Generic binding protein
7	<u>AKAP12</u>	A-kinase anchor protein 12	Generic binding protein
8	<u>AMN1</u>	Protein AMN1 homolog	Generic binding protein
9	<u>AMPD3</u>	AMP deaminase 3	Generic enzyme
10	<u>ANKRD37</u>	Ankyrin repeat domain-containing protein 37	Generic binding protein
11	<u>ANKRD50</u>	Ankyrin repeat domain-containing protein 50	Generic protein
12	<u>AP1S1</u>	AP-1 complex subunit sigma-1A	Transporter
13	<u>APLP2</u>	Amyloid-like protein 2	Generic binding protein
14	<u>ARG2</u>	Arginase-2, mitochondrial	Generic enzyme
15	<u>ARHGAP18</u>	Rho GTPase-activating protein 18	Regulators (GDI, GAP, GEF etc.)
16	<u>ARHGAP22</u>	Rho GTPase-activating protein 22	Regulators (GDI, GAP, GEF etc.)
17	<u>ARHGAP5</u>	Rho GTPase-activating protein 5	Regulators (GDI, GAP, GEF etc.)
18	<u>ARL1</u>	ADP-ribosylation factor-like protein 1	RAS superfamily
19	<u>ASMTL</u>	N-acetylserotonin O-methyltransferase-like protein	Generic enzyme
20	<u>ATM</u>	Serine-protein kinase ATM	Protein kinase
21	<u>ATP13A5</u>	Probable cation-transporting ATPase 13A5	Generic binding protein
22	<u>ATP1A1</u>	Sodium/potassium-transporting ATPase subunit alpha-1	Generic channel
23	<u>ATP1B1</u>	Sodium/potassium-transporting ATPase subunit beta-1	Generic binding protein
24	<u>ATP2C1</u>	Calcium-transporting ATPase type 2C member 1	Generic channel
25	<u>B4GALT7</u>	Beta-1,4-galactosyltransferase 7	Generic enzyme
26	<u>BAT2L2</u>	BAT2 domain-containing protein 1	Generic protein
27	<u>BBS2</u>	Bardet-Biedl syndrome 2 protein	Generic binding protein
28	<u>BNC2</u>	Zinc finger protein basoonuclin-2	Generic binding protein
29	<u>C13orf36</u>	Uncharacterized protein C13orf36	Generic protein
30	<u>C13orf37</u>	UPF0582 protein C13orf37	Generic protein
31	<u>C18orf19</u>	Uncharacterized protein C18orf19	Generic protein
32	<u>C18orf32</u>	UPF0729 protein C18orf32	Generic protein
33	<u>C20orf199</u>		RNA
34	<u>C21orf91</u>	Protein EURL homolog	Generic protein
35	<u>C6orf62</u>	Uncharacterized protein C6orf62	Generic protein
36	<u>CALR</u>	Calreticulin	Generic binding protein

37	<u>CAMK2D</u>	Calcium/calmodulin-dependent protein kinase type II delta chain	Protein kinase
38	<u>CAV2</u>	Caveolin-2	Generic binding protein
39	<u>CCNG1</u>	Cyclin-G1	Generic binding protein
40	<u>CD274</u>	Programmed cell death 1 ligand 1	Generic binding protein
41	<u>CD302</u>	CD302 antigen	Generic binding protein
42	<u>CD44</u>	CD44 antigen	Generic receptor
43	<u>CD46</u>	Membrane cofactor protein	Generic receptor
44	<u>CDC42SE2</u>	CDC42 small effector protein 2	Generic binding protein
45	<u>CEP57</u>	Centrosomal protein of 57 kDa	Generic binding protein
46	<u>CEP63</u>	Centrosomal protein of 63 kDa	Generic binding protein
47	<u>CLASP1</u>	CLIP-associating protein 1	Generic binding protein
48	<u>CLEC11A</u>	C-type lectin domain family 11 member A	Receptor ligand
49	<u>CLK4</u>	Dual specificity protein kinase CLK4	Protein kinase
50	<u>CNN3</u>	Calponin-3	Generic binding protein
51	<u>CNOT1</u>	CCR4-NOT transcription complex subunit 1	Generic binding protein
52	<u>COG6</u>	Conserved oligomeric Golgi complex subunit 6	Generic binding protein
53	<u>CPD</u>	Carboxypeptidase D	Generic protease
54	<u>CRTAP</u>	Cartilage-associated protein	Generic binding protein
55	<u>CUL3</u>	Cullin-3	Generic binding protein
56	<u>DCBLD2</u>	Discoidin, CUB and LCCL domain-containing protein 2	Generic binding protein
57	<u>DCUN1D4</u>	DCN1-like protein 4	Generic protein
58	<u>DDX17</u>	Probable ATP-dependent RNA helicase DDX17	Generic enzyme
59	<u>DFNA5</u>	Non-syndromic hearing impairment protein 5	Generic binding protein
60	<u>DRAM2</u>	DNA damage-regulated autophagy modulator protein 2	Generic protein
61	<u>DSE</u>	Dermatan-sulfate epimerase	Generic enzyme
62	<u>EEF1A1</u>	Elongation factor 1-alpha 1	Generic binding protein
63	<u>EEF1B2</u>	Elongation factor 1-beta	Generic binding protein
64	<u>EEF1B4</u>		
65	<u>EGFR</u>	Epidermal growth factor receptor	Receptor with enzyme activity
66	<u>EIF2S2</u>	Eukaryotic translation initiation factor 2 subunit 2	Generic binding protein
67	<u>EIF3E</u>	Eukaryotic translation initiation factor 3 subunit E	Generic binding protein
68	<u>EIF4B</u>	Eukaryotic translation initiation factor 4B	Generic binding protein
69	<u>ELL2</u>	RNA polymerase II elongation factor ELL2	Generic binding protein
70	<u>EPRS</u>	Bifunctional aminoacyl-tRNA synthetase	Generic enzyme
71	<u>ERAP1</u>	Endoplasmic reticulum aminopeptidase 1	Metalloprotease
72	<u>ERN1</u>	Serine/threonine-protein kinase/endoribonuclease IRE1	Protein kinase
73	<u>ERO1L</u>	ERO1-like protein alpha	Generic enzyme
74	<u>ERP44</u>	Endoplasmic reticulum resident protein ERp44	Generic binding protein
75	<u>ETV3</u>	ETS translocation variant 3	Transcription factor

76	<u>EXD3</u>	Probable exonuclease mut-7 homolog, isoform 5	Generic enzyme
		Probable exonuclease mut-7 homolog	Generic binding protein
77	<u>FAM13A</u>	Protein FAM13A	Generic protein
78	<u>FAM162A</u>	UPF0389 protein FAM162A	Generic protein
79	<u>FAM167A</u>	UPF0484 protein FAM167A	Generic protein
80	<u>FAM43A</u>	Protein FAM43A	Generic protein
81	<u>FN1</u>	Fibronectin	Receptor ligand
82	<u>FNDC3A</u>	Fibronectin type-III domain-containing protein 3a	Generic protein
83	<u>FRG1</u>	Protein FRG1	Transporter
84	<u>FRG1B</u>	Protein FRG1B	Generic protein
85	<u>FUBP1</u>	Far upstream element-binding protein 1	Transcription factor
86	<u>FXR1</u>	Fragile X mental retardation syndrome-related protein 1	Generic binding protein
87	<u>GALNT12</u>	Polypeptide N-acetylgalactosaminyltransferase 12	Generic enzyme
88	<u>GALNT5</u>	Polypeptide N-acetylgalactosaminyltransferase 5	Generic enzyme
89	<u>GAS5</u>		RNA
90	<u>GLS</u>	Glutaminase kidney isoform, mitochondrial	Generic enzyme
91	<u>GNA13</u>	Guanine nucleotide-binding protein subunit alpha-13	G-alpha
92	<u>GOLGB1</u>	Golgin subfamily B member 1	Generic binding protein
93	<u>GUSB</u>	Beta-glucuronidase	Generic enzyme
94	<u>HEY1</u>	Hairy/enhancer-of-split related with YRPW motif protein 1	Transcription factor
95	<u>HLTF</u>	Helicase-like transcription factor	Generic binding protein
96	<u>HNRNPD</u>	Heterogeneous nuclear ribonucleoprotein D0	Generic binding protein
97	<u>HNRPDL</u>	Heterogeneous nuclear ribonucleoprotein D-like	Generic binding protein
98	<u>HOMER1</u>	Homer protein homolog 1	Generic binding protein
99	<u>HSP90B1</u>	Endoplasmic	Generic binding protein
100	<u>HSPA2</u>	Heat shock-related 70 kDa protein 2	Generic binding protein
101	<u>HSPA5</u>	78 kDa glucose-regulated protein	Generic binding protein
102	<u>IAH1</u>	Isoamyl acetate-hydrolyzing esterase 1 homolog	Generic enzyme
103	<u>IER3IP1</u>	Immediate early response 3-interacting protein 1	Generic binding protein
104	<u>IER5L</u>	Immediate early response gene 5-like protein	Generic protein
105	<u>IGFBP3</u>	Insulin-like growth factor-binding protein 3	Generic binding protein
106	<u>IL17RB</u>	Interleukin-17 receptor B	Generic receptor
107	<u>IL8</u>	Interleukin-8	Receptor ligand
108	<u>INSIG2</u>	Insulin-induced gene 2 protein	Generic binding protein
109	<u>INTS6</u>	Integrator complex subunit 6	Generic enzyme
110	<u>IRS2</u>	Insulin receptor substrate 2	Generic binding protein
111	<u>ITGA4</u>	Integrin alpha-4	Generic receptor
112	<u>ITGB1</u>	Integrin beta-1	Generic receptor

113	<u>IVNS1ABP</u>	Influenza virus NS1A-binding protein	Generic binding protein
114	<u>JARID2</u>	Protein Jumonji	Generic binding protein
115	<u>KCNMA1</u>	Calcium-activated potassium channel subunit alpha-1	Voltage-gated ion-channel
116	<u>KDM5B</u>	Lysine-specific demethylase 5B	Transcription factor
117	<u>KIAA1143</u>	Uncharacterized protein KIAA1143	Generic protein
118	<u>KIAA1244</u>	Brefeldin A-inhibited guanine nucleotide-exchange protein 3	Regulators (GDI, GAP, GEF etc.)
119	<u>KIAA1245</u>	KIAA1245	Generic protein
120	<u>KISS1R</u>	KiSS-1 receptor	GPCR
121	<u>LAPTM4A</u>	Lysosomal-associated transmembrane protein 4A	Generic protein
122	<u>LAYN</u>	Layilin	Generic receptor
123	<u>LCORL</u>	Ligand-dependent nuclear receptor corepressor-like protein	Generic binding protein
124	<u>LEMD1</u>	LEM domain-containing protein 1	Generic protein
125	<u>LEPR</u>	Leptin receptor	Generic receptor
126	<u>LGALS8</u>	Galectin-8	Receptor ligand
127	<u>LIPG</u>	Endothelial lipase	Generic phospholipase
128	<u>LMAN1</u>	Protein ERGIC-53	Generic binding protein
129	<u>LOC100133315</u>	Putative short transient receptor potential channel 2-like protein	Generic binding protein
130	<u>LOC100134091</u>		
131	<u>LOC149832</u>		
132	<u>LOC151162</u>		
133	<u>LOC387763</u>	Protein Ag2 homolog	Generic protein
134	<u>LOC399959</u>		
135	<u>LOC440957</u>	UPF0640 protein	Generic binding protein
136	<u>LOC550643</u>	Putative uncharacterized protein LOC550643	Generic protein
137	<u>LOC642236</u>		
138	<u>LOC643837</u>		RNA
139	<u>LOC653380</u>	TBC1 domain family member 3C-like protein ENSP00000341742	Regulators (GDI, GAP, GEF etc.)
140	<u>LOC727820</u>		
141	<u>LOC729082</u>		
142	<u>LOC729222</u>		
143	<u>LOC730510</u>		
144	<u>LONP2</u>	Peroxisomal Lon protease homolog 2	Generic protease
145	<u>MAN2A1</u>	Alpha-mannosidase 2	Generic enzyme
146	<u>MAP3K2</u>	Mitogen-activated protein kinase kinase kinase 2	Protein kinase
147	<u>MAP4K4</u>	Mitogen-activated protein kinase kinase kinase kinase 4	Protein kinase
148	<u>MAPK6</u>	Mitogen-activated protein kinase 6	Protein kinase
149	<u>MBNL2</u>	Muscleblind-like protein 2	Generic binding protein
150	<u>MED21</u>	Mediator of RNA polymerase II transcription subunit 21	Generic binding protein



151	<u>MIB1</u>	E3 ubiquitin-protein ligase MIB1	Generic enzyme
152	<u>MIR21</u>		RNA
153	<u>MLH3</u>	DNA mismatch repair protein Mlh3	Generic binding protein
154	<u>MPHOSPH6</u>	M-phase phosphoprotein 6	Generic binding protein
155	<u>MPZL2</u>	Myelin protein zero-like protein 2	Generic binding protein
156	<u>MRI1</u>	Methylthioribose-1-phosphate isomerase	Generic binding protein
157	<u>MT1A</u>	Metallothionein-1A	Generic binding protein
158	<u>MT1E</u>	Metallothionein-1E	Generic binding protein
159	<u>MT1H</u>	Metallothionein-1H	Generic binding protein
160	<u>MT1M</u>	Metallothionein-1M	Generic binding protein
161	<u>MT1P2</u>	MT-1H-like protein	Generic protein
162	<u>MXI1</u>	MAX-interacting protein 1	Transcription factor
163	<u>N4BP2L2</u>	NEDD4-binding protein 2-like 2	Generic protein
164	<u>NACA</u>	Nascent polypeptide-associated complex subunit alpha	Generic binding protein
165	<u>NAP1L1</u>	Nucleosome assembly protein 1-like 1	Generic binding protein
166	<u>NBAS</u>	Neuroblastoma-amplified sequence	Generic protein
167	<u>NBPF1</u>	Neuroblastoma breakpoint family member 1	Generic protein
168	<u>NBPF10</u>	Neuroblastoma breakpoint family member 10	Generic protein
169	<u>NBPF11</u>	Neuroblastoma breakpoint family member 11	Generic protein
170	<u>NBPF14</u>	Neuroblastoma breakpoint family member 14	Generic protein
171	<u>NBPF15</u>	Neuroblastoma breakpoint family member 15	Generic protein
172	<u>NBPF16</u>	Neuroblastoma breakpoint family member 16	Generic protein
173	<u>NBPF3</u>	Neuroblastoma breakpoint family member 3	Generic protein
174	<u>NBPF8</u>	Neuroblastoma breakpoint family member 21	Generic protein
		Neuroblastoma breakpoint family member 8	Generic protein
175	<u>NBPF9</u>	Neuroblastoma breakpoint family member 20	Generic protein
176	<u>NCRNA00219</u>	Putative uncharacterized protein TIGA1	Generic protein
177	<u>NDP</u>	Norrin	Receptor ligand
178	<u>NDUFB5</u>	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 5, mitochondrial	Generic enzyme
179	<u>NDUFS4</u>	NADH dehydrogenase [ubiquinone] iron-sulfur protein 4, mitochondrial	Generic enzyme
180	<u>NFIB</u>	Nuclear factor 1 B-type	Transcription factor
181	<u>NFXL1</u>	NF-X1-type zinc finger protein NFXL1	Generic binding protein
182	<u>NGLY1</u>	Peptide-N(4)-(N-acetyl-beta-glucosaminyl)asparagine amidase	Generic enzyme
183	<u>NMD3</u>	60S ribosomal export protein NMD3	Generic binding protein
184	<u>NOL3</u>	Nucleolar protein 3	Generic binding protein
185	<u>NOL7</u>	Nucleolar protein 7	Generic protein
186	<u>NOTCH2NL</u>	Notch homolog 2 N-terminal-like protein	Generic binding protein
187	<u>NT5E</u>	5'-nucleotidase	Generic phosphatase
188	<u>NUCB2</u>	Nucleobindin-2	Generic binding protein
189	<u>OCIAD1</u>	OCIA domain-containing protein 1	Generic protein
190	<u>P4HA2</u>	Prolyl 4-hydroxylase subunit alpha-2	Generic enzyme
191	<u>PABPC1L</u>	Polyadenylate-binding protein 1-like	Generic binding protein

192	<u>PAM</u>	Peptidyl-glycine alpha-amidating monooxygenase	Generic enzyme
193	<u>PAQR5</u>	Membrane progesterin receptor gamma	Generic receptor
194	<u>PCNP</u>	PEST proteolytic signal-containing nuclear protein	Generic binding protein
195	<u>PCNX</u>	Pecanex-like protein 1	Generic protein
196	<u>PDCD6</u>	Programmed cell death protein 6	Generic binding protein
197	<u>PDGFC</u>	Platelet-derived growth factor C	Receptor ligand
198	<u>PDIA6</u>	Protein disulfide-isomerase A6	Generic enzyme
199	<u>PECI</u>	Peroxisomal 3,2-trans-enoyl-CoA isomerase	Generic enzyme
200	<u>PER3</u>	Period circadian protein homolog 3	Generic binding protein
201	<u>PGCP</u>	Plasma glutamate carboxypeptidase	Generic protease
202	<u>PHF20</u>	PHD finger protein 20	Generic binding protein
203	<u>PLCG2</u>	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma-2	Generic phospholipase
204	<u>PLEKHA3</u>	Pleckstrin homology domain-containing family A member 3	Generic binding protein
205	<u>PLLP</u>	Plasmolipin	Generic channel
206	<u>PMPCB</u>	Mitochondrial-processing peptidase subunit beta	Metalloprotease
207	<u>PNN</u>	Pinin	Transcription factor
208	<u>PPARA</u>	Peroxisome proliferator-activated receptor alpha	Transcription factor
209	<u>PPFIBP1</u>	Liprin-beta-1	Generic binding protein
210	<u>PPIL3</u>	Peptidyl-prolyl cis-trans isomerase-like 3	Generic enzyme
211	<u>PPP1R9A</u>	Neurabin-1	Generic binding protein
212	<u>PRKRA</u>	Interferon-inducible double stranded RNA-dependent protein kinase activator A	Generic binding protein
213	<u>PRUNE2</u>	Protein prune homolog 2	Generic protein
		BNIP2 motif-containing molecule at the C-terminal region 1	Generic protein
214	<u>PTEN</u>	Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN	Lipid phosphatase
215	<u>PTENP1</u>	phosphatase and tensin homolog (mutated in multiple advanced cancers 1), pseudogene 1	Protein phosphatase
216	<u>PURA</u>	Transcriptional activator protein Pur-alpha	Transcription factor
217	<u>PYROXD1</u>	Pyridine nucleotide-disulfide oxidoreductase domain-containing protein 1	Generic enzyme
218	<u>QPCT</u>	Glutamyl-peptide cyclotransferase	Generic enzyme
219	<u>RAB11FIP3</u>	Rab11 family-interacting protein 3	Regulators (GDI, GAP, GEF etc.)
220	<u>RAB13</u>	Ras-related protein Rab-13	RAS superfamily
221	<u>RAP2B</u>	Ras-related protein Rap-2b	RAS superfamily
222	<u>RARS2</u>	Probable arginyl-tRNA synthetase, mitochondrial	Generic enzyme
223	<u>RBM12</u>	RNA-binding protein 12	Generic binding protein
224	<u>RBM5</u>	RNA-binding protein 5	Generic binding protein
225	<u>RBM6</u>	RNA-binding protein 6	Generic binding protein

226	<u>RBPJ</u>	Recombining binding protein suppressor of hairless	Transcription factor
227	<u>RGS10</u>	Regulator of G-protein signaling 10	Regulators (GDI, GAP, GEF etc.)
228	<u>RGS5</u>	Regulator of G-protein signaling 5	Regulators (GDI, GAP, GEF etc.)
229	<u>RIMKLB</u>	Ribosomal protein S6 modification-like protein B	Generic enzyme
230	<u>RNASE4</u>	Ribonuclease 4	Generic enzyme
231	<u>RNF160</u>	RING finger protein 160	Generic binding protein
232	<u>RP11-94I2.2</u>	neuroblastoma breakpoint family, member 11-like	Generic protein
233	<u>RPL14</u>	60S ribosomal protein L14	Generic binding protein
234	<u>RPL14P1</u>		
235	<u>RPL17</u>	60S ribosomal protein L17	Generic binding protein
236	<u>RPL17P6</u>		
237	<u>RPL22</u>	60S ribosomal protein L22	Generic binding protein
238	<u>RPL3</u>	60S ribosomal protein L3	Generic binding protein
239	<u>RPL30</u>	60S ribosomal protein L30	Generic binding protein
240	<u>RPL31</u>	60S ribosomal protein L31	Generic binding protein
241	<u>RPL36A</u>	60S ribosomal protein L36a	Generic binding protein
242	<u>RPL36AP37</u>	similar to large subunit ribosomal protein L36a	Generic protein
243	<u>RPL9</u>	60S ribosomal protein L9	Generic binding protein
244	<u>RPS10</u>	40S ribosomal protein S10	Generic binding protein
245	<u>RPS15A</u>	40S ribosomal protein S15a	Generic binding protein
246	<u>RPS15AP19</u>		
247	<u>RPS7</u>	40S ribosomal protein S7	Generic binding protein
248	<u>RSBN1</u>	Round spermatid basic protein 1	Generic binding protein
249	<u>RTTN</u>	Rotatin	Generic binding protein
250	<u>S100A10</u>	Protein S100-A10	Receptor ligand
251	<u>SAT2</u>	Diamine acetyltransferase 2	Generic enzyme
252	<u>SC65</u>	Synaptonemal complex protein SC65	Generic protein
253	<u>SCAMP1</u>	Secretory carrier-associated membrane protein 1	Generic binding protein
254	<u>SCD</u>	Putative uncharacterized protein PRO1933	Generic protein
		Acyl-CoA desaturase	Generic enzyme
255	<u>SDC2</u>	Syndecan-2	Generic receptor
256	<u>SEC22B</u>	Vesicle-trafficking protein SEC22b	Generic binding protein
257	<u>SEC62</u>	Translocation protein SEC62	Generic receptor
258	<u>SEL1L</u>	Protein sel-1 homolog 1	Generic binding protein
259	<u>SELK</u>	Selenoprotein K	Generic binding protein
260	<u>SEPT8</u>	Septin-8	Generic binding protein
261	<u>SFRS11</u>	Splicing factor, arginine/serine-rich 11	Generic binding protein
262	<u>SGK3</u>	Serine/threonine-protein kinase Sgk3	Protein kinase
263	<u>SH3KBP1</u>	SH3 domain-containing kinase-binding protein 1	Generic binding protein
264	<u>SLAMF7</u>	SLAM family member 7	Generic receptor

265	<u>SLC25A36</u>	Solute carrier family 25 member 36	Transporter
266	<u>SLC2A3</u>	Solute carrier family 2, facilitated glucose transporter member 3	Transporter
267	<u>SLC5A3</u>	Sodium/myo-inositol cotransporter	Transporter
268	<u>SMAD9</u>	Mothers against decapentaplegic homolog 9	Transcription factor
269	<u>SNHG5</u>		RNA
270	<u>SNHG6</u>		
271	<u>SNHG8</u>		
272	<u>SNORD50A</u>		RNA
273	<u>SNORD50B</u>		
274	<u>SNX21</u>	Sorting nexin-21	Generic binding protein
275	<u>SNX3</u>	Sorting nexin-3	Generic binding protein
276	<u>SNX5</u>	Sorting nexin-5	Generic binding protein
277	<u>SOX4</u>	Transcription factor SOX-4	Transcription factor
278	<u>SPAG4</u>	Sperm-associated antigen 4 protein	Generic binding protein
279	<u>SRGAP2P1</u>	SLIT-ROBO Rho GTPase activating protein 2 pseudogene 1	Generic protein
280	<u>SSR1</u>	Translocon-associated protein subunit alpha	Generic binding protein
281	<u>ST3GAL6</u>	Type 2 lactosamine alpha-2,3-sialyltransferase	Generic enzyme
282	<u>SYNE2</u>	Nesprin-2	Generic binding protein
283	<u>SYTL2</u>	Synaptotagmin-like protein 2	Generic binding protein
284	<u>TAF1D</u>	TATA box-binding protein-associated factor RNA polymerase I subunit D	Generic binding protein
285	<u>TALDO1</u>	Transaldolase	Generic enzyme
286	<u>TAX1BP1</u>	Tax1-binding protein 1	Generic binding protein
287	<u>TBC1D3</u>	TBC1 domain family member 3E/3F	Generic protein
288	<u>TBC1D3C</u>	TBC1 domain family member 3C/3D	Regulators (GDI, GAP, GEF etc.)
289	<u>TBC1D3F</u>	TBC1 domain family member 3	Regulators (GDI, GAP, GEF etc.)
290	<u>TBC1D3G</u>	TBC1 domain family member 3G	Regulators (GDI, GAP, GEF etc.)
291	<u>TBC1D3H</u>	TBC1 domain family member 3-like protein LOC729837	Regulators (GDI, GAP, GEF etc.)
		TBC1 domain family member 3E/3F	Generic protein
292	<u>TFB1M</u>	Dimethyladenosine transferase 1, mitochondrial	Generic enzyme
293	<u>THBD</u>	Thrombomodulin	Generic receptor
294	<u>TKT</u>	Transketolase	Generic enzyme
295	<u>TM4SF1</u>	Transmembrane 4 L6 family member 1	Generic receptor
296	<u>TM4SF4</u>	Transmembrane 4 L6 family member 4	Generic receptor
297	<u>TM9SF3</u>	Transmembrane 9 superfamily member 3	Generic binding protein
298	<u>TMED10</u>	Transmembrane emp24 domain-containing protein 10	Generic binding protein
299	<u>TMEM45B</u>	Transmembrane protein 45B	Generic protein
300	<u>TMEM49</u>	Transmembrane protein 49	Generic protein
301	<u>TMEM9</u>	Transmembrane protein 9	Generic protein

302	<u>TMTC2</u>	Transmembrane and TPR repeat-containing protein 2	Generic binding protein
303	<u>TPP1</u>	Tripeptidyl-peptidase 1	Generic protease
304	<u>TRPM8</u>	Transient receptor potential cation channel subfamily M member 8	Generic channel
305	<u>TTC3</u>	Tetratricopeptide repeat protein 3	Generic binding protein
306	<u>TXNIP</u>	Thioredoxin-interacting protein	Generic binding protein
307	<u>UBE2I</u>	SUMO-conjugating enzyme UBC9	Generic enzyme
308	<u>USO1</u>	General vesicular transport factor p115	Transporter
309	<u>VAMP4</u>	Vesicle-associated membrane protein 4	Generic protein
310	<u>VEGFA</u>	Vascular endothelial growth factor A	Receptor ligand
311	<u>VKORC1</u>	Vitamin K epoxide reductase complex subunit 1	Generic enzyme
312	<u>VLDLR</u>	Very low-density lipoprotein receptor	Generic receptor
313	<u>WDR66</u>	WD repeat-containing protein 66	Generic binding protein
314	<u>WSB1</u>	WD repeat and SOCS box-containing protein 1	Generic binding protein
315	<u>XIST</u>		RNA
316	<u>XYLT2</u>	Xylosyltransferase 2	Generic enzyme
317	<u>YME1L1</u>	ATP-dependent metalloprotease YME1L1	Metalloprotease
318	<u>ZAK</u>	Mitogen-activated protein kinase kinase kinase MLT	Protein kinase
319	<u>ZMYM2</u>	Zinc finger MYM-type protein 2	Generic binding protein
320	<u>ZNF12</u>	Zinc finger protein 12	Generic binding protein
321	<u>ZNF57</u>	Zinc finger protein 57	Generic binding protein
322	<u>ZNF804A</u>	Zinc finger protein 804A	Generic binding protein
323	<u>ZNF84</u>	Zinc finger protein 84	Generic binding protein
324	<u>ZNF91</u>	Zinc finger protein 91	Transcription factor
325	<u>hCG 19809</u>		

**Table 2.0 T-test cluster 2**

#	Gene Symbol	Protein name	Object type
1	<u>ABCF2</u>	ATP-binding cassette sub-family F member 2	Transporter
2	<u>ACAP3</u>	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 3	Regulators (GDI, GAP, GEF etc.)
3	<u>ACIN1</u>	Apoptotic chromatin condensation inducer in the nucleus	Generic binding protein
4	<u>ACSL5</u>	Long-chain-fatty-acid--CoA ligase 5	Generic enzyme
5	<u>ACTBP9</u>		
6	<u>ADAMTSL4</u>	ADAMTS-like protein 4	Metalloprotease
7	<u>ADRM1</u>	Proteasomal ubiquitin receptor ADRM1	Generic binding protein
8	<u>AFG3L2</u>	AFG3-like protein 2	Metalloprotease
9	<u>AGRN</u>	Agrin	Generic binding protein
10	<u>AKIRIN2</u>	Akirin-2	Generic binding protein
11	<u>AKT1</u>	RAC-alpha serine/threonine-protein kinase	Protein kinase
12	<u>ANKRD40</u>	Ankyrin repeat domain-containing protein 40	Generic protein
13	<u>ANXA7</u>	Annexin A7	Generic binding protein
14	<u>AP1M1</u>	AP-1 complex subunit mu-1	Transporter
15	<u>APOA1BP</u>	Apolipoprotein A-I-binding protein	Generic binding protein
16	<u>APRT</u>	Adenine phosphoribosyltransferase	Generic enzyme
17	<u>ARF3</u>	ADP-ribosylation factor 3	RAS superfamily
18	<u>ARID1A</u>	AT-rich interactive domain-containing protein 1A	Generic binding protein
19	<u>ARID3A</u>	AT-rich interactive domain-containing protein 3A	Transcription factor
20	<u>ARL2BP</u>	ADP-ribosylation factor-like protein 2-binding protein	Generic binding protein
21	<u>ARPC4</u>	Actin-related protein 2/3 complex subunit 4	Generic binding protein
22	<u>ATF6</u>	Cyclic AMP-dependent transcription factor ATF-6 alpha	Transcription factor
23	<u>ATP6V0C</u>	V-type proton ATPase 16 kDa proteolipid subunit	Generic protein
24	<u>ATP6V0D1</u>	V-type proton ATPase subunit d 1	Transporter
25	<u>ATXN1L</u>	Ataxin-1-like	Generic binding protein
26	<u>BAG5</u>	BAG family molecular chaperone regulator 5	Generic binding protein
27	<u>BAT1</u>	Spliceosome RNA helicase BAT1	Generic enzyme
28	<u>BRMS1</u>	Breast cancer metastasis-suppressor 1	Generic binding protein
29	<u>C18orf10</u>	Tubulin polyglutamylase complex subunit 2	Generic protein
30	<u>C19orf21</u>	Uncharacterized protein C19orf21	Generic protein
31	<u>C1orf85</u>	Lysosomal protein NCU-G1	Generic protein
32	<u>C20orf108</u>	Transmembrane protein C20orf108	Generic protein
33	<u>C20orf29</u>	Uncharacterized protein C20orf29	Generic protein
34	<u>C2orf29</u>	Uncharacterized protein C2orf29	Generic protein
35	<u>C7orf49</u>	Modulator of retrovirus infection homolog	Generic protein
36	<u>C9orf167</u>	Torsin family protein C9orf167	Generic binding protein
37	<u>CABIN1</u>	Calcineurin-binding protein cabin-1	Generic binding protein

38	<u>CALM1</u>	Calmodulin	Generic binding protein
39	<u>CALM2</u>	Calmodulin	Generic binding protein
40	<u>CALM3</u>	Calmodulin	Generic binding protein
41	<u>CAPRIN1</u>	Caprin-1	Generic binding protein
42	<u>CAPZB</u>	F-actin-capping protein subunit beta	Generic binding protein
43	<u>CBS</u>	Cystathionine beta-synthase	Generic enzyme
44	<u>CC2D1A</u>	Coiled-coil and C2 domain-containing protein 1A	Generic binding protein
45	<u>CCDC109A</u>	Coiled-coil domain-containing protein 109A	Generic protein
46	<u>CCDC85C</u>	Coiled-coil domain-containing protein 85C	Generic protein
47	<u>CCDC90A</u>	Coiled-coil domain-containing protein 90A, mitochondrial	Generic protein
48	<u>CCDC97</u>	Coiled-coil domain-containing protein 97	Generic protein
49	<u>CDS2</u>	Phosphatidate cytidyltransferase 2	Generic enzyme
50	<u>CEACAM1</u>	Carcinoembryonic antigen-related cell adhesion molecule 1	Generic protein
51	<u>CENPB</u>	Major centromere autoantigen B	Generic binding protein
52	<u>CISH</u>	Cytokine-inducible SH2-containing protein	Generic binding protein
53	<u>CLN3</u>	Battenin	Generic binding protein
54	<u>CLPTM1</u>	Cleft lip and palate transmembrane protein 1	Generic protein
55	<u>CNN3</u>	Calponin-3	Generic binding protein
56	<u>CNP</u>	2',3'-cyclic-nucleotide 3'-phosphodiesterase	Generic enzyme
57	<u>COMMD4</u>	COMM domain-containing protein 4	Generic binding protein
58	<u>COX5A</u>	Cytochrome c oxidase subunit 5A, mitochondrial	Generic enzyme
59	<u>CSNK1E</u>	Casein kinase I isoform epsilon	Protein kinase
60	<u>CTBP1</u>	C-terminal-binding protein 1	Generic binding protein
61	<u>CTDSP1</u>	Carboxy-terminal domain RNA polymerase II polypeptide A small phosphatase 1	Protein phosphatase
62	<u>CYB5B</u>	Cytochrome b5 type B	Transporter
63	<u>DCTN1</u>	Dynactin subunit 1	Generic binding protein
64	<u>DDX28</u>	Probable ATP-dependent RNA helicase DDX28	Generic enzyme
65	<u>DDX56</u>	Probable ATP-dependent RNA helicase DDX56	Generic enzyme
66	<u>DIDO1</u>	Death-inducer obliterator 1	Generic binding protein
67	<u>DMWD</u>	Dystrophia myotonica WD repeat-containing protein	Generic protein
68	<u>DNAJB6</u>	DnaJ homolog subfamily B member 6	Generic binding protein
69	<u>DNAJC14</u>	DnaJ homolog subfamily C member 14	Generic binding protein
70	<u>DULLARD</u>	Serine/threonine-protein phosphatase dullard	Generic phosphatase
71	<u>E2F4</u>	Transcription factor E2F4	Transcription factor
72	<u>EAPP</u>	E2F-associated phosphoprotein	Generic protein
73	<u>EIF3B</u>	Eukaryotic translation initiation factor 3 subunit B	Generic binding protein
74	<u>EIF4G1</u>	Eukaryotic translation initiation factor 4 gamma 1	Generic binding protein
75	<u>ERBB2</u>	Receptor tyrosine-protein kinase erbB-2	Receptor with enzyme activity

76	<u>ESRRA</u>	Steroid hormone receptor ERR1	Transcription factor
77	<u>EVC</u>	Ellis-van Creveld syndrome protein	Generic protein
78	<u>EVI5L</u>	EVI5-like protein	Regulators (GDI, GAP, GEF etc.)
79	<u>EWSR1</u>	RNA-binding protein EWS	Generic binding protein
80	<u>EXOC6B</u>	Exocyst complex component 6B	Generic protein
81	<u>EXOC7</u>	Exocyst complex component 7	Generic protein
82	<u>FAM100A</u>	Protein FAM100A	Generic protein
83	<u>FAM21A</u>	Protein FAM21A	Generic phosphatase
84	<u>FAM21B</u>	Protein FAM21B	Generic protein
85	<u>FAM21C</u>	Protein FAM21C	Generic phosphatase
86	<u>FAM21D</u>	Protein FAM21D	Generic protein
87	<u>FAM38A</u>	Protein FAM38A	Generic protein
88	<u>FAM53C</u>	Protein FAM53C	Generic protein
89	<u>FAM86B1</u>	Protein FAM86B1	Generic protein
90	<u>FARS2</u>	Phenylalanyl-tRNA synthetase, mitochondrial	Generic enzyme
91	<u>FBXL14</u>	F-box/LRR-repeat protein 14	Generic binding protein
92	<u>FBXO7</u>	F-box only protein 7	Generic binding protein
93	<u>FKBP1A</u>	Peptidyl-prolyl cis-trans isomerase FKBP1A	Generic binding protein
94	<u>FKRP</u>	Fukutin-related protein	Generic enzyme
95	<u>FLNA</u>	Filamin-A	Generic binding protein
96	<u>FUT1</u>	Galactoside 2-alpha-L-fucosyltransferase 1	Generic enzyme
97	<u>G3BP1</u>	Ras GTPase-activating protein-binding protein 1	Generic enzyme
98	<u>GAMT</u>	Guanidinoacetate N-methyltransferase	Generic enzyme
99	<u>GAS2L1</u>	GAS2-like protein 1	Generic protein
100	<u>GBA</u>	Glucosylceramidase	Generic enzyme
101	<u>GBAP</u>		
102	<u>GCN1L1</u>	Translational activator GCN1	Generic binding protein
103	<u>GNA11</u>	Guanine nucleotide-binding protein subunit alpha-11	G-alpha
104	<u>GNAI2</u>	Guanine nucleotide-binding protein G(i), alpha-2 subunit	G-alpha
105	<u>GOLGA6L4</u>	Putative golgin subfamily A member 6-like protein 4	Generic protein
106	<u>GPS2</u>	G protein pathway suppressor 2	Regulators (GDI, GAP, GEF etc.)
107	<u>GRINA</u>	Glutamate [NMDA] receptor-associated protein 1	Ligand-gated ion-channel
108	<u>GSPT1</u>	Eukaryotic peptide chain release factor GTP-binding subunit ERF3A	Generic binding protein
109	<u>GTF2F1</u>	General transcription factor IIF subunit 1	Transcription factor
110	<u>H1FX</u>	Histone H1x	Generic binding protein
111	<u>H2AFV</u>	Histone H2A.V	Generic binding protein
112	<u>HDGF</u>	Hepatoma-derived growth factor	Generic binding protein
113	<u>HEXIM1</u>	Protein HEXIM1	Generic binding protein
114	<u>HIST1H2BC</u>	Histone H2B type 1-C/E/F/G/I	Generic binding protein



115	<u>HIST1H2BE</u>	Histone H2B type 1-C/E/F/G/I	Generic binding protein
116	<u>HIST1H2BF</u>	Histone H2B type 1-C/E/F/G/I	Generic binding protein
117	<u>HIST1H2BG</u>	Histone H2B type 1-C/E/F/G/I	Generic binding protein
118	<u>HIST1H2BI</u>	Histone H2B type 1-C/E/F/G/I	Generic binding protein
119	<u>HMG20A</u>	High mobility group protein 20A	Transcription factor
120	<u>HMGB3</u>	High mobility group protein B3	Generic binding protein
121	<u>HN1</u>	Hematological and neurological expressed 1 protein	Generic protein
122	<u>HNRNPH2</u>	Heterogeneous nuclear ribonucleoprotein H2	Generic binding protein
123	<u>HNRNPU</u>	Heterogeneous nuclear ribonucleoprotein U	Generic binding protein
124	<u>ICOSLG</u>	ICOS ligand	Generic receptor
125	<u>ICT1</u>	Immature colon carcinoma transcript 1 protein	Generic binding protein
126	<u>IFIT5</u>	Interferon-induced protein with tetratricopeptide repeats 5	Generic binding protein
127	<u>IFNGR2</u>	Interferon-gamma receptor beta chain	Generic receptor
128	<u>ILF2</u>	Interleukin enhancer-binding factor 2	Transcription factor
129	<u>INO80E</u>	INO80 complex subunit E	Generic protein
130	<u>INTS5</u>	Integrator complex subunit 5	Generic binding protein
131	<u>IPO4</u>	Importin-4	Transporter
132	<u>IRF9</u>	Interferon regulatory factor 9	Transcription factor
133	<u>ISY1</u>	Pre-mRNA-splicing factor ISY1 homolog	Generic protein
134	<u>KDM3B</u>	Lysine-specific demethylase 3B	Generic enzyme
135	<u>KIAA0196</u>	Strumpellin	Generic protein
136	<u>KIAA0226</u>	Uncharacterized protein KIAA0226	Generic protein
137	<u>KIAA0284</u>	Protein KIAA0284	Generic protein
138	<u>KIAA1279</u>	KIF1-binding protein	Generic binding protein
139	<u>KIAA1737</u>	Uncharacterized protein KIAA1737	Generic protein
140	<u>KLF13</u>	Krueppel-like factor 13	Transcription factor
141	<u>KRT7</u>	Keratin, type II cytoskeletal 7	Generic binding protein
142	<u>LARP1</u>	La-related protein 1	Generic binding protein
143	<u>LASP1</u>	LIM and SH3 domain protein 1	Transporter
144	<u>LDOC1L</u>	Protein LDOC1L	Generic protein
145	<u>LEMD2</u>	LEM domain-containing protein 2	Generic protein
146	<u>LIMK2</u>	LIM domain kinase 2	Protein kinase
147	<u>LOC100128612</u>		
148	<u>LOC150776</u>	Putative uncharacterized protein FLJ41352	Generic protein
149	<u>LSM12</u>	Protein LSM12 homolog	Generic binding protein
150	<u>MAD1L1</u>	Mitotic spindle assembly checkpoint protein MAD1	Generic binding protein
151	<u>MAP4</u>	Microtubule-associated protein 4	Generic binding protein
152	<u>MARK2</u>	Serine/threonine-protein kinase MARK2	Protein kinase
153	<u>MARVELD1</u>	Putative MARVEL domain-containing protein 1	Generic protein
154	<u>MAT2A</u>	S-adenosylmethionine synthetase isoform type-2	Generic enzyme
155	<u>MCOLN1</u>	Mucolipin-1	Generic channel

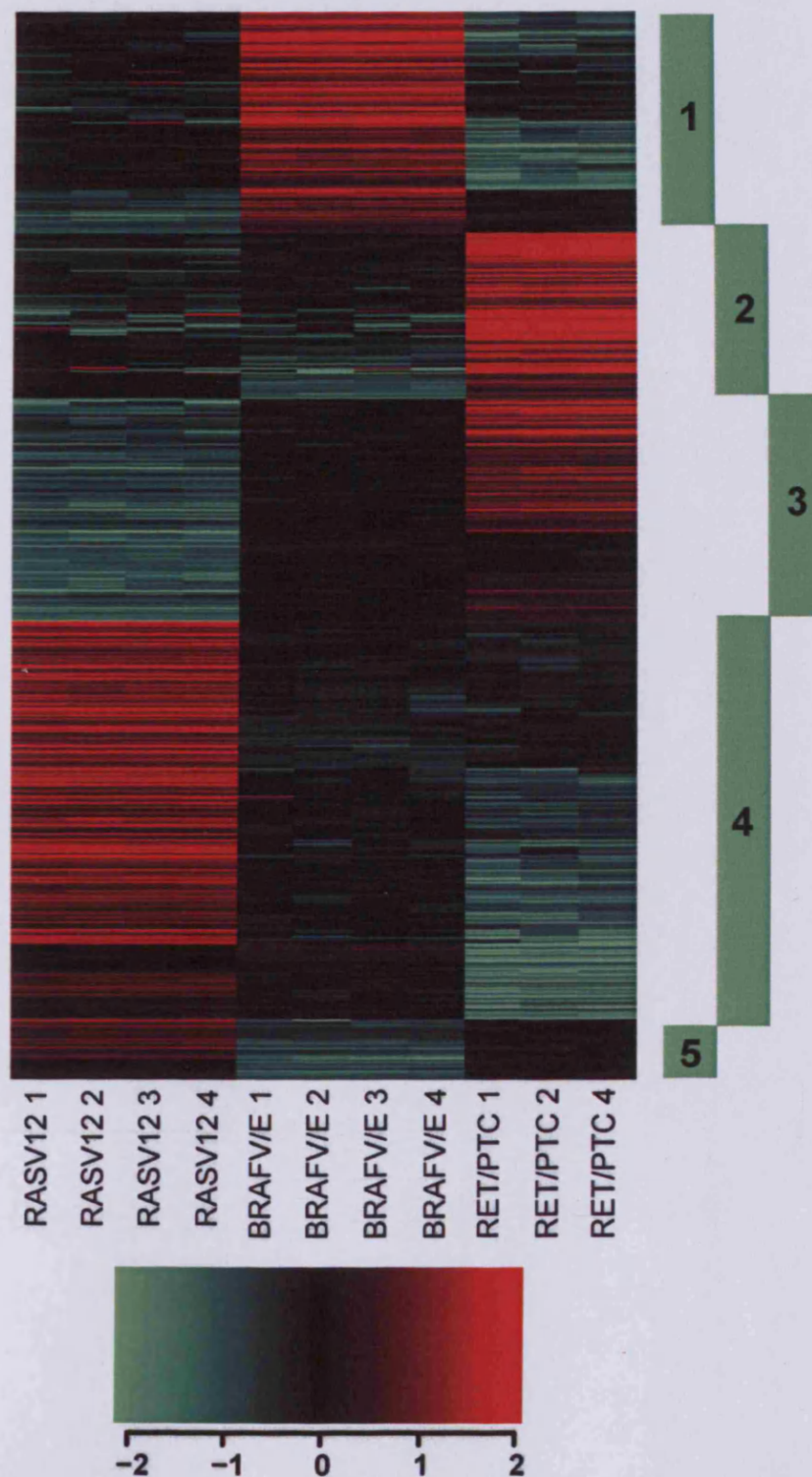
156	<u>MED22</u>	Mediator of RNA polymerase II transcription subunit 22	Generic binding protein
157	<u>METTL13</u>	Methyltransferase-like protein 13	Generic enzyme
158	<u>MFN2</u>	Mitofusin-2	Generic binding protein
159	<u>MICALL1</u>	MICAL-like protein 1	Generic binding protein
160	<u>MIIP</u>	Migration and invasion-inhibitory protein	Generic protein
161	<u>MINK1</u>	Misshapen-like kinase 1	Protein kinase
162	<u>MKL1</u>	MKL/myocardin-like protein 1	Generic binding protein
163	<u>MLLT6</u>	Protein AF-17	Generic binding protein
164	<u>MON1A</u>	Vacuolar fusion protein MON1 homolog A	Generic protein
165	<u>MPDU1</u>	Mannose-P-dolichol utilization defect 1 protein	Generic binding protein
166	<u>MPRIIP</u>	Myosin phosphatase Rho-interacting protein	Generic binding protein
167	<u>MSL1</u>	Male-specific lethal 1 homolog	Generic binding protein
168	<u>MYO18A</u>	Myosin-XVIIIa	Generic binding protein
169	<u>NACC1</u>	Nucleus accumbens-associated protein 1	Generic binding protein
170	<u>NADK</u>	NAD kinase	Generic kinase
171	<u>NAPA</u>	Alpha-soluble NSF attachment protein	Transporter
172	<u>NBL1</u>	Neuroblastoma suppressor of tumorigenicity 1	Generic binding protein
173	<u>NCOA5</u>	Nuclear receptor coactivator 5	Generic binding protein
174	<u>NECAP1</u>	Adaptin ear-binding coat-associated protein 1	Generic binding protein
175	<u>NKIRAS2</u>	NF-kappa-B inhibitor-interacting Ras-like protein 2	RAS superfamily
176	<u>NOC4L</u>	Nucleolar complex protein 4 homolog	Generic binding protein
177	<u>NUP50</u>	Nucleoporin 50 kDa	Generic channel
178	<u>NUP62</u>	Nuclear pore glycoprotein p62	Generic channel
179	<u>OPTN</u>	Optineurin	Generic binding protein
180	<u>OR7E104P</u>		
181	<u>PACS2</u>	Phosphofurin acidic cluster sorting protein 2	Generic protein
182	<u>PAIP1</u>	Polyadenylate-binding protein-interacting protein 1	Generic binding protein
183	<u>PARP9</u>	Poly [ADP-ribose] polymerase 9	Generic enzyme
184	<u>PATZ1</u>	POZ-, AT hook-, and zinc finger-containing protein 1	Generic binding protein
185	<u>PBRM1</u>	Protein polybromo-1	Generic binding protein
186	<u>PCNXL3</u>	Pecanex-like protein 3	Generic protein
187	<u>PCYT1A</u>	Choline-phosphate cytidyltransferase A	Generic enzyme
188	<u>PEBP1</u>	Phosphatidylethanolamine-binding protein 1	Generic binding protein
189	<u>PEF1</u>	Peflin	Generic binding protein
190	<u>PFN1</u>	Profilin-1	Generic binding protein
191	<u>PGAP2</u>	Post-GPI attachment to proteins factor 2	Generic binding protein
192	<u>PGLS</u>	6-phosphogluconolactonase	Generic enzyme
193	<u>PGRMC2</u>	Membrane-associated progesterone receptor component 2	Generic receptor
194	<u>PHF12</u>	PHD finger protein 12	Generic binding protein
195	<u>PHLDA3</u>	Pleckstrin homology-like domain family A member 3	Generic binding protein

196	<u>PI4KA</u>	Phosphatidylinositol 4-kinase alpha	Lipid kinase
197	<u>PI4KAP1</u>	Putative phosphatidylinositol 4-kinase alpha-like protein P1	Lipid kinase
198	<u>PI4KAP2</u>	Putative phosphatidylinositol 4-kinase alpha-like protein P2	Generic protein
199	<u>PIP5K1C</u>	Phosphatidylinositol-4-phosphate 5-kinase type-1 gamma	Generic kinase
200	<u>PISD</u>	Phosphatidylserine decarboxylase proenzyme	Generic enzyme
201	<u>PITPNA</u>	Phosphatidylinositol transfer protein alpha isoform	Transporter
202	<u>PLOD3</u>	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	Generic enzyme
203	<u>PMEPA1</u>	Transmembrane prostate androgen-induced protein	Generic binding protein
204	<u>PML</u>	Probable transcription factor PML	Transcription factor
205	<u>PNMA1</u>	Paraneoplastic antigen Ma1	Generic binding protein
206	<u>PNPO</u>	Pyridoxine-5'-phosphate oxidase	Generic enzyme
207	<u>POLDIP2</u>	Polymerase delta-interacting protein 2	Generic binding protein
208	<u>POLRMT</u>	DNA-directed RNA polymerase, mitochondrial	Generic enzyme
209	<u>POM121</u>	Nuclear envelope pore membrane protein POM 121	Generic channel
210	<u>POM121B</u>	Putative nuclear envelope pore membrane protein POM 121B	Generic protein
211	<u>POM121C</u>	Nuclear envelope pore membrane protein POM 121C	Generic channel
212	<u>PPM1K</u>	Protein phosphatase 1K, mitochondrial	Protein phosphatase
213	<u>PPP1R10</u>	Serine/threonine-protein phosphatase 1 regulatory subunit 10	Generic binding protein
214	<u>PPP1R9B</u>	Neurabin-2	Generic binding protein
215	<u>PPP4R1</u>	Serine/threonine-protein phosphatase 4 regulatory subunit 1	Generic binding protein
216	<u>PPP5C</u>	Serine/threonine-protein phosphatase 5	Protein phosphatase
217	<u>PRCC</u>	Proline-rich protein PRCC	Generic protein
218	<u>PREB</u>	Prolactin regulatory element-binding protein	Generic binding protein
219	<u>PRIC285</u>	Peroxisomal proliferator-activated receptor A-interacting complex 285 kDa protein	Generic binding protein
220	<u>PRICKLE1</u>	Prickle-like protein 1	Generic binding protein
221	<u>PRKD2</u>	Serine/threonine-protein kinase D2	Protein kinase
222	<u>PSMB2</u>	Proteasome subunit beta type-2	Generic protease
223	<u>PSMB6</u>	Proteasome subunit beta type-6	Generic protease
224	<u>PSMD3</u>	26S proteasome non-ATPase regulatory subunit 3	Generic binding protein
225	<u>PSME3</u>	Proteasome activator complex subunit 3	Generic binding protein
226	<u>PTBP1</u>	Polypyrimidine tract-binding protein 1	Generic binding protein
227	<u>PTPN9</u>	Tyrosine-protein phosphatase non-receptor type 9	Protein phosphatase
228	<u>PTPRJ</u>	Receptor-type tyrosine-protein phosphatase eta	Generic receptor
229	<u>PYGO2</u>	Pygopus homolog 2	Generic binding protein

230	<u>RANBP10</u>	Ran-binding protein 10	Generic protein
231	<u>RAVER1</u>	Ribonucleoprotein PTB-binding 1	Generic binding protein
232	<u>RBBP4</u>	Histone-binding protein RBBP4	Generic enzyme
233	<u>RBM12</u>	RNA-binding protein 12	Generic binding protein
234	<u>RBMS1</u>	RNA-binding motif, single-stranded-interacting protein 1	Generic binding protein
235	<u>RCE1</u>	CAAX prenyl protease 2	Generic protease
236	<u>RHOG</u>	Rho-related GTP-binding protein RhoG	RAS superfamily
237	<u>RNF144B</u>	E3 ubiquitin-protein ligase RNF144B	Generic enzyme
238	<u>RNF19B</u>	E3 ubiquitin-protein ligase RNF19B	Generic binding protein
239	<u>RNF220</u>	RING finger protein 220	Generic binding protein
240	<u>RNF31</u>	RING finger protein 31	Generic binding protein
241	<u>RNF40</u>	E3 ubiquitin-protein ligase BRE1B	Generic enzyme
242	<u>RNFT1</u>	RING finger and transmembrane domain-containing protein 1	Generic binding protein
243	<u>RRBP1</u>	Ribosome-binding protein 1	Generic receptor
244	<u>S100A2</u>	Protein S100-A2	Generic binding protein
245	<u>SAE1</u>	SUMO-activating enzyme subunit 1	Generic enzyme
246	<u>SARNP</u>	SAP domain-containing ribonucleoprotein	Generic binding protein
247	<u>SBF1</u>	Myotubularin-related protein 5	Protein phosphatase
248	<u>SENP3</u>	Sentrin-specific protease 3	Generic protease
249	<u>SF1</u>	Splicing factor 1	Generic binding protein
250	<u>SFRS1</u>	Splicing factor, arginine/serine-rich 1	Generic binding protein
251	<u>SFRS15</u>	Splicing factor, arginine/serine-rich 15	Generic binding protein
252	<u>SFRS2B</u>	Splicing factor, arginine/serine-rich 2B	Generic binding protein
253	<u>SFT2D2</u>	Vesicle transport protein SFT2B	Generic protein
254	<u>SGPL1</u>	Sphingosine-1-phosphate lyase 1	Generic enzyme
255	<u>SGSH</u>	N-sulphoglucosamine sulphohydrolase	Generic enzyme
256	<u>SH2D3A</u>	SH2 domain-containing protein 3A	Generic binding protein
257	<u>SIK3</u>	Serine/threonine-protein kinase SIK3	Protein kinase
258	<u>SIRT6</u>	NAD-dependent deacetylase sirtuin-6	Generic enzyme
259	<u>SIVA1</u>	Apoptosis regulatory protein Siva	Generic binding protein
260	<u>SLC35C1</u>	GDP-fucose transporter 1	Transporter
261	<u>SLC4A2</u>	Anion exchange protein 2	Transporter
262	<u>SLC4A5</u>	Electrogenic sodium bicarbonate cotransporter 4	Transporter
263	<u>SMARCA4</u>	Probable global transcription activator SNF2L4	Generic enzyme
264	<u>SMG5</u>	Protein SMG5	Generic binding protein
265	<u>SMG7</u>	Protein SMG7	Generic binding protein
266	<u>SMPD4</u>	Sphingomyelin phosphodiesterase 4	Generic enzyme
267	<u>SPATA24</u>	Testis protein T6441 homolog	Generic protein
268	<u>SPHK2</u>	Sphingosine kinase 2	Lipid kinase
269	<u>SPOP</u>	Speckle-type POZ protein	Generic binding protein
270	<u>SPTAN1</u>	Spectrin alpha chain, brain	Generic binding protein
271	<u>SRI</u>	Sorcin	Generic binding protein

272	<u>STAT3</u>	Signal transducer and activator of transcription 3	Transcription factor
273	<u>STAT6</u>	Signal transducer and activator of transcription 6	Transcription factor
274	<u>STOML1</u>	Stomatin-like protein 1	Generic binding protein
275	<u>STRN4</u>	Striatin-4	Generic binding protein
276	<u>STUB1</u>	STIP1 homology and U box-containing protein 1	Generic enzyme
277	<u>SUPT6H</u>	Transcription elongation factor SPT6	Transcription factor
278	<u>SYNCRIP</u>	Heterogeneous nuclear ribonucleoprotein Q	Generic binding protein
279	<u>TAGLN2</u>	Transgelin-2	Generic protein
280	<u>TBC1D25</u>	TBC1 domain family member 25	Regulators (GDI, GAP, GEF etc.)
281	<u>TBCC</u>	Tubulin-specific chaperone C	Generic binding protein
282	<u>TCEB3</u>	Transcription elongation factor B polypeptide 3	Generic binding protein
283	<u>TENC1</u>	Tensin-like C1 domain-containing phosphatase	Generic binding protein
284	<u>TEX261</u>	Protein TEX261	Generic protein
285	<u>TFE3</u>	Transcription factor E3	Transcription factor
286	<u>TGOLN2</u>	Trans-Golgi network integral membrane protein 2	Generic binding protein
287	<u>TH1L</u>	Negative elongation factor C/D	Generic binding protein
288	<u>TIAF1</u>	TGFB1-induced anti-apoptotic factor 1	Generic binding protein
289	<u>TIMM23</u>	Mitochondrial import inner membrane translocase subunit Tim23	Generic protein
290	<u>TMEM139</u>	Transmembrane protein 139	Generic protein
291	<u>TMEM160</u>	Transmembrane protein 160	Generic protein
292	<u>TMEM185B</u>	Transmembrane protein 185B	Generic protein
293	<u>TMEM199</u>	Transmembrane protein 199	Generic protein
294	<u>TMEM201</u>	Transmembrane protein 201	Generic protein
295	<u>TNRC18</u>	Trinucleotide repeat-containing gene 18 protein	Generic binding protein
296	<u>TNRC6C</u>	Trinucleotide repeat-containing gene 6C protein	Generic binding protein
297	<u>TOPORS</u>	E3 ubiquitin-protein ligase Topors	Generic enzyme
298	<u>TRAF4</u>	TNF receptor-associated factor 4	Generic binding protein
299	<u>TRIM34</u>	Tripartite motif-containing protein 34	Generic binding protein
300	<u>TRIM6</u>	Tripartite motif-containing protein 6	Generic binding protein
301	<u>TRIM6-TRIM34</u>	interferon-responsive finger protein 1 long form	Generic protein
302	<u>TRIM65</u>	Tripartite motif-containing protein 65	Generic binding protein
303	<u>TRIM69</u>	Tripartite motif-containing protein 69	Generic binding protein
304	<u>TSPAN14</u>	Tetraspanin-14	Generic receptor
305	<u>TSPYL5</u>	Testis-specific Y-encoded-like protein 5	Generic protein
306	<u>TSSC4</u>	Protein TSSC4	Generic protein
307	<u>TSTA3</u>	GDP-L-fucose synthetase	Generic enzyme
308	<u>TTLL3</u>	Tubulin monoglycylase TTLL3	Generic enzyme
309	<u>UBA1</u>	Ubiquitin-like modifier-activating enzyme 1	Generic enzyme
310	<u>UBAP2L</u>	Ubiquitin-associated protein 2-like	Generic binding protein
311	<u>UBE2A</u>	Ubiquitin-conjugating enzyme E2 A	Generic enzyme

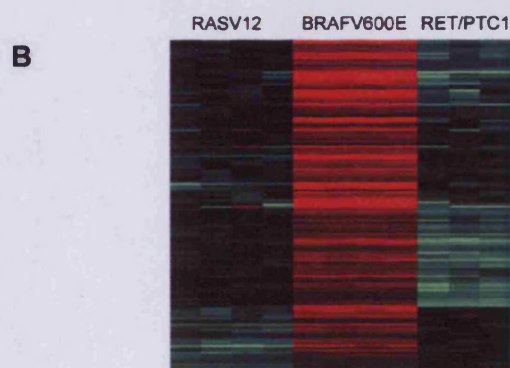
312	<u>UBE2D2</u>	Ubiquitin-conjugating enzyme E2 D2	Generic enzyme
313	<u>UBE2M</u>	NEDD8-conjugating enzyme Ubc12	Generic enzyme
314	<u>UBE2N</u>	Ubiquitin-conjugating enzyme E2 N	Generic enzyme
315	<u>UBE3C</u>	Ubiquitin-protein ligase E3C	Generic enzyme
316	<u>UMPS</u>	Uridine 5'-monophosphate synthase	Generic enzyme
317	<u>UQCRC1</u>	Cytochrome b-c1 complex subunit 1, mitochondrial	Generic binding protein
318	<u>URM1</u>	Ubiquitin-related modifier 1 homolog	Generic protein
319	<u>VAPB</u>	Vesicle-associated membrane protein-associated protein B/C	Generic binding protein
320	<u>VDR</u>	Vitamin D3 receptor	Transcription factor
321	<u>WDR18</u>	WD repeat-containing protein 18	Generic protein
322	<u>WDR5</u>	WD repeat-containing protein 5	Generic binding protein
323	<u>WDR77</u>	Methylosome protein 50	Generic binding protein
324	<u>WIZ</u>	Protein Wiz	Generic binding protein
325	<u>WWC1</u>	Protein WWC1	Generic binding protein
326	<u>YWHAB</u>	14-3-3 protein beta/alpha	Generic binding protein
327	<u>YWHAH</u>	14-3-3 protein eta	Generic binding protein
328	<u>ZDHHC12</u>	Probable palmitoyltransferase ZDHHC12	Generic enzyme
329	<u>ZDHHC3</u>	Palmitoyltransferase ZDHHC3	Generic enzyme
330	<u>ZDHHC8</u>	Probable palmitoyltransferase ZDHHC8	Generic enzyme
331	<u>ZFAND3</u>	AN1-type zinc finger protein 3	Generic binding protein
332	<u>ZNF532</u>	Zinc finger protein 532	Generic binding protein
333	<u>ZNF592</u>	Zinc finger protein 592	Generic binding protein
334	<u>ZNF593</u>	Zinc finger protein 593	Generic binding protein



Used correlation distance metric.  
average linkage clustering, on logcen data.

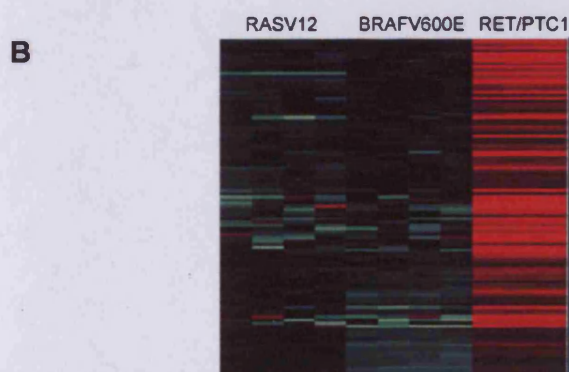
**Figure 4.** ANOVA of genotype-specific changes in gene expression. Hierarchical clustering displaying the genotype specific transcriptional changes of 690 differentially regulated probesets in RASV12, BRAF<sup>V600E</sup> and RET/PTC1 oncogene infected thyrocytes. The columns correspond to the oncogene and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation.



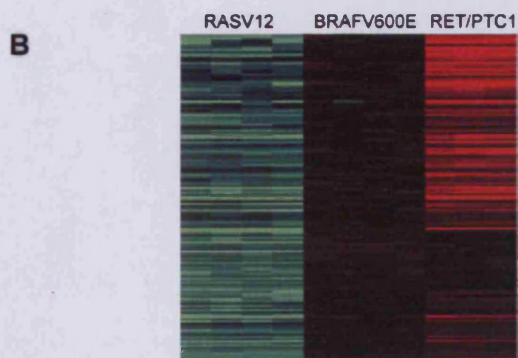


**Figure 5.** Enrichment analysis of pathway folders identified in ANOVA cluster 1. The enrichment method was subjected to an FDR of 0.05. **A**, The order of map folders is based on their significance in the cluster. The dark orange bars dictate significance set at 0.05%. Although the light/orange bars are displayed they are not discussed due to the possibility of false positive results. For the ANOVA cluster 1, there were 21 map folders identified with 1 of significance. **B**, The corresponding heatmap of the ANOVA cluster 1, displaying an overall expression pattern of 121 median/down-regulated probes in RASV12, down-regulated probes in RET/PTC1 and high/up-regulation in BRAF<sup>V600E</sup>. The columns correspond to the oncogenic condition and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation, whilst median expression is signified by black.



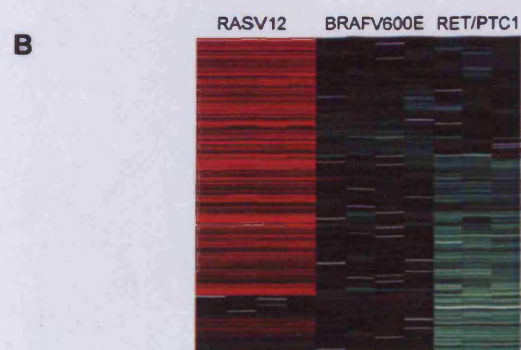


**Figure 6.** Enrichment analysis of pathway folders identified in ANOVA cluster 2. The enrichment method was subjected to an FDR of 0.05. **A**, The order of map folders is based on their significance in the cluster. The dark orange bars dictate significance set at 0.05%. Although the light/orange bars are displayed they are not discussed due to the possibility of false positive results. For the ANOVA cluster 2, there were 32 map folders identified with 1 of significance. **B**, The corresponding heatmap of the ANOVA cluster 2, displaying an overall expression pattern of 96 median/down-regulated probes in RASV12, and BRAF<sup>V600E</sup> and high/up-regulated probes in RET/PTC1. The columns correspond to the oncogenic condition and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation, whilst median expression is signified by black.

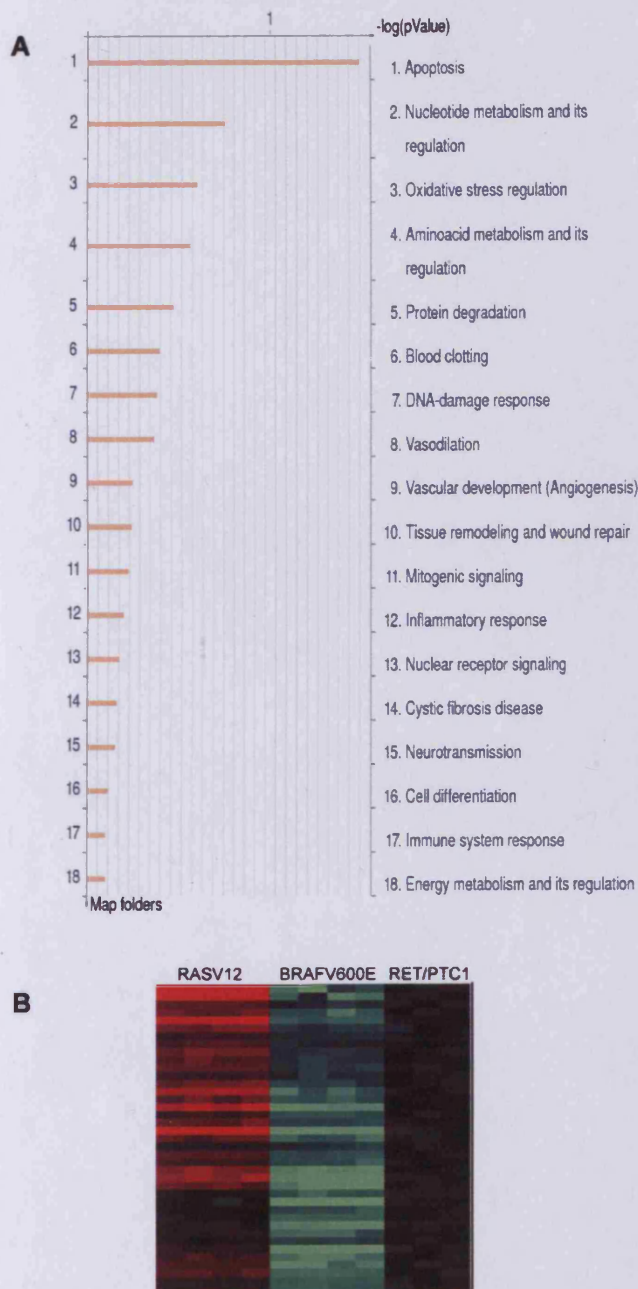


**Figure 7.** Enrichment analysis of pathway folders identified in ANOVA cluster 3. The enrichment method was subjected to an FDR of 0.05. **A**, The order of map folders is based on their significance in the cluster. The dark orange bars dictate significance set at 0.05%. Although the light/orange bars are displayed they are not discussed due to the possibility of false positive results. For the ANOVA cluster 3, there were 32 map folders identified with 9 of significance. **B**, The corresponding heatmap of the ANOVA cluster 3, displaying an overall expression pattern of 142 up-regulated/highly expressed probes in RET/PTC1, and downregulated/low expression in RASV12 and median expression in BRAF<sup>V600E</sup>. The columns correspond to the oncogenic condition and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation, whilst median expression is signified by black.





**Figure 8.** Enrichment analysis of pathway folders identified in ANOVA cluster 4. The enrichment method was subjected to an FDR of 0.05. **A**, The order of map folders is based on their significance in the cluster. The dark orange bars dictate significance set at 0.05%. Although the light/orange bars are displayed they are not discussed due to the possibility of false positive results. For the ANOVA cluster 4, there were 32 map folders identified with 5 of significance. **B**, The corresponding heatmap of the ANOVA cluster 4, displaying an overall expression pattern of 209 high/up-regulated probes in RASV12, median expression in BRAF<sup>V600E</sup> and median/low down-regulated probes in RET/PTC1. The columns correspond to the oncogenic condition and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation, whilst median expression is signified by black.



**Figure 9.** Enrichment analysis of pathway folders identified in ANOVA cluster 5. The enrichment method was subjected to an FDR of 0.05. **A**, The order of map folders is based on their significance in the cluster. The dark orange bars dictate significance set at 0.05%. Although the light/orange bars are displayed they are not discussed due to the possibility of false positive results. For the ANOVA cluster 5, there were 18 map folders identified with no significant maps. **B**, The corresponding heatmap of the ANOVA cluster 5, displaying an overall expression pattern of 34 high/up-regulated probes in RASV12, median expression in RET/PTC1 and low/down-regulated probes in BRAF<sup>V600E</sup>. The columns correspond to the oncogenic condition and the rows to each gene transcript. Red signifies up-regulation and green signifies down-regulation, whilst median expression is signified by black.

**Table 3.0 ANOVA Cluster 1**

#	Gene Symbol	Protein name	Object type
1	<u>ABL2</u>	Tyrosine-protein kinase ABL2	Protein kinase
2	<u>ADAMTS18</u>	A disintegrin and metalloproteinase with thrombospondin motifs 18	Metalloprotease
3	<u>AJAP1</u>	Adherens junction-associated protein 1	Generic binding protein
4	<u>ARID3A</u>	AT-rich interactive domain-containing protein 3A	Transcription factor
5	<u>BOK</u>	Bcl-2-related ovarian killer protein	Generic binding protein
6	<u>C1orf38</u>	Protein THEMIS2	Generic enzyme
7	<u>C2orf54</u>	Uncharacterized protein C2orf54	Generic protein
8	<u>CA2</u>	Carbonic anhydrase 2	Generic enzyme
9	<u>CALCOCO1</u>	Calcium-binding and coiled-coil domain-containing protein 1	Generic binding protein
10	<u>CD9</u>	CD9 antigen	Receptor ligand
11	<u>CDC42EP1</u>	Cdc42 effector protein 1	Generic binding protein
12	<u>CDSN</u>	Corneodesmosin	Generic binding protein
13	<u>CDYL</u>	Chromodomain Y-like protein	Generic binding protein
14	<u>CGN</u>	Cingulin	Generic binding protein
15	<u>CGNL1</u>	Cingulin-like protein 1	Generic binding protein
16	<u>CLCF1</u>	Cardiotrophin-like cytokine factor 1	Receptor ligand
17	<u>CLDN6</u>	Claudin-6	Generic binding protein
18	<u>CLSTN2</u>	Calsyntenin-2	Generic binding protein
19	<u>CLTB</u>	Clathrin light chain B	Generic binding protein
20	<u>CNKSR1</u>	Connector enhancer of kinase suppressor of ras 1	Generic binding protein
21	<u>CORO1B</u>	Coronin-1B	Generic binding protein
22	<u>CPPED1</u>	Calcineurin-like phosphoesterase domain-containing protein 1	Generic enzyme
23	<u>CYP4F11</u>	Cytochrome P450 4F11	Generic enzyme
24	<u>DBNDD2</u>	Dysbindin domain-containing protein 2	Generic binding protein
25	<u>DHRS9</u>	Dehydrogenase/reductase SDR family member 9	Generic enzyme
26	<u>DIRC2</u>	Disrupted in renal carcinoma protein 2	Generic protein
27	<u>DOPEY2</u>	Protein dopey-2	Generic protein
28	<u>DUSP10</u>	Dual specificity protein phosphatase 10	Protein phosphatase
29	<u>DUSP5</u>	Dual specificity protein phosphatase 5	Protein phosphatase
30	<u>ELF3</u>	ETS-related transcription factor Elf-3	Transcription factor
31	<u>EPS8L1</u>	Epidermal growth factor receptor kinase substrate 8-like protein 1	Generic binding protein

32	<u>ERBB3</u>	Receptor tyrosine-protein kinase erbB-3	Receptor with enzyme activity
33	<u>ESAM</u>	Endothelial cell-selective adhesion molecule	Generic binding protein
34	<u>F11R</u>	Junctional adhesion molecule A	Generic binding protein
35	<u>FAM150B</u>	Protein FAM150B	Generic protein
36	<u>FAM83A</u>	Protein FAM83A	Generic protein
37	<u>FAM83H</u>	Protein FAM83H	Generic protein
38	<u>FAM84B</u>	Protein FAM84B	Generic binding protein
39	<u>GNAL</u>	Guanine nucleotide-binding protein G(olf) subunit alpha	G-alpha
40	<u>HCG22</u>		
41	<u>HDGF</u>	Hepatoma-derived growth factor	Generic binding protein
42	<u>HOMER2</u>	Homer protein homolog 2	Generic binding protein
43	<u>HPGD</u>	15-hydroxyprostaglandin dehydrogenase [NAD+]	Generic enzyme
44	<u>HSD17B2</u>	Estradiol 17-beta-dehydrogenase 2	Generic enzyme
45	<u>IL1RN</u>	Interleukin-1 receptor antagonist protein	Receptor ligand
46	<u>ITGA2</u>	Integrin alpha-2	Generic receptor
47	<u>ITPKC</u>	Inositol-trisphosphate 3-kinase C	Generic kinase
48	<u>IVL</u>	Involucrin	Generic binding protein
49	<u>JUP</u>	Junction plakoglobin	Generic binding protein
50	<u>KCNF1</u>	Potassium voltage-gated channel subfamily F member 1	Voltage-gated ion-channel
51	<u>KIF1A</u>	Kinesin-like protein KIF1A	Generic binding protein
52	<u>KLF6</u>	Krueppel-like factor 6	Transcription factor
53	<u>KLHDC3</u>	Kelch domain-containing protein 3	Generic binding protein
54	<u>KLK10</u>	Kallikrein-10	Generic protease
55	<u>KLK6</u>	Kallikrein-6	Generic protease
56	<u>KLK7</u>	Kallikrein-7	Generic protease
57	<u>L1CAM</u>	Neural cell adhesion molecule L1	Generic binding protein
58	<u>LAMC2</u>	Laminin subunit gamma-2	Receptor ligand
59	<u>LCN2</u>	Neutrophil gelatinase-associated lipocalin	Transporter
60	<u>LPIN1</u>	Lipin-1	Generic binding protein
61	<u>LYPLA1</u>	Acyl-protein thioesterase 1	Generic enzyme
62	<u>MEAF6</u>	Chromatin modification-related protein MEAF6	Generic protein
63	<u>MED15</u>	Mediator of RNA polymerase II transcription subunit 15	Generic binding protein
64	<u>METRNL</u>	Meteorin	Generic protein
65	<u>MFI2</u>	Melanotransferrin	Generic binding protein
66	<u>MMP10</u>	Stromelysin-2	Metalloprotease
67	<u>MUC20</u>	Mucin-20	Generic binding protein

68	<u>MYCN</u>	N-myc proto-oncogene protein	Transcription factor
69	<u>MYO1C</u>	Myosin-Ic	Generic binding protein
70	<u>MYO1D</u>	Myosin-IId	Generic binding protein
71	<u>NDFIP1</u>	NEDD4 family-interacting protein 1	Generic binding protein
72	<u>NDRG4</u>	Protein NDRG4	Generic protein
73	<u>PHYHIPL</u>	Phytanoyl-CoA hydroxylase-interacting protein-like	Generic protein
74	<u>PI3</u>	Elafin	Generic binding protein
75	<u>PLA2G2F</u>	Group IIF secretory phospholipase A2	Generic phospholipase
76	<u>PLEKHA7</u>	Pleckstrin homology domain-containing family A member 7	Generic binding protein
77	<u>PPL</u>	Periplakin	Generic binding protein
78	<u>PRPS1</u>	Ribose-phosphate pyrophosphokinase 1	Generic kinase
79	<u>PRSS22</u>	Brain-specific serine protease 4	Generic protease
80	<u>PRSS37</u>	Probable inactive trypsin-X2	Generic protein
81	<u>PRSS8</u>	Prostasin	Generic protease
82	<u>PSG1</u>	Pregnancy-specific beta-1-glycoprotein 1	Generic protein
83	<u>PSG4</u>	Pregnancy-specific beta-1-glycoprotein 4	Generic protein
84	<u>PSG9</u>	Pregnancy-specific beta-1-glycoprotein 9	Generic protein
85	<u>RAB11FIP1</u>	Rab11 family-interacting protein 1	Generic binding protein
86	<u>RAB15</u>	Ras-related protein Rab-15	RAS superfamily
87	<u>RALBP1</u>	RalA-binding protein 1	Regulators (GDI, GAP, GEF etc.)
88	<u>RNF14</u>	E3 ubiquitin-protein ligase RNF14	Generic enzyme
89	<u>RNF39</u>	RING finger protein 39	Generic binding protein
90	<u>RP1-21O18.1</u>	Kazrin	Generic protein
91	<u>RRAS2</u>	Ras-related protein R-Ras2	RAS superfamily
92	<u>RTN2</u>	Reticulon-2	Generic binding protein
93	<u>SCNN1A</u>	Amiloride-sensitive sodium channel subunit alpha	Generic protein
94	<u>SDCBP2</u>	Syntenin-2	Generic binding protein
95	<u>SEC14L2</u>	SEC14-like protein 2	Transporter
96	<u>SERPINB1</u>	Leukocyte elastase inhibitor	Generic binding protein
97	<u>SIRPA</u>	Tyrosine-protein phosphatase non-receptor type substrate 1	Generic binding protein
98	<u>SLC12A6</u>	Solute carrier family 12 member 6	Transporter
99	<u>SLC4A11</u>	Sodium bicarbonate transporter-like protein 11	Transporter
100	<u>SMURF1</u>	E3 ubiquitin-protein ligase SMURF1	Generic enzyme
101	<u>SMURF2</u>	E3 ubiquitin-protein ligase SMURF2	Generic enzyme
102	<u>SOX17</u>	Transcription factor SOX-17	Transcription factor
103	<u>SPRR1A</u>	Cornifin-A	Generic binding protein

104	<u>SPRR1B</u>	Cornifin-B	Generic binding protein
105	<u>ST8SIA4</u>	CMP-N-acetylneuraminate-poly-alpha-2,8-sialyltransferase	Generic enzyme
106	<u>SULT2B1</u>	Sulfotransferase family cytosolic 2B member 1	Generic enzyme
107	<u>SVIL</u>	Supervillin	Generic binding protein
108	<u>SYS1-DBNDD2</u>		
109	<u>THSD7A</u>	Thrombospondin type-1 domain-containing protein 7A	Generic protein
110	<u>TICAM1</u>	TIR domain-containing adapter molecule 1	Generic binding protein
111	<u>TJP3</u>	Tight junction protein ZO-3	Generic binding protein
112	<u>TM7SF2</u>	Delta(14)-sterol reductase	Generic enzyme
113	<u>TMCC3</u>	Transmembrane and coiled-coil domains protein 3	Generic protein
114	<u>TSHZ1</u>	Teashirt homolog 1	Generic binding protein
115	<u>TTC9</u>	Tetratricopeptide repeat protein 9A	Generic binding protein
116	<u>UCA1</u>		
117	<u>UXS1</u>	UDP-glucuronic acid decarboxylase 1	Generic enzyme
118	<u>WDR45</u>	WD repeat domain phosphoinositide-interacting protein 4	Generic protein
119	<u>WNT9A</u>	Protein Wnt-9a	Receptor ligand
120	<u>ZADH2</u>	Zinc-binding alcohol dehydrogenase domain-containing protein 2	Generic enzyme
121	<u>ZDHHC13</u>	Probable palmitoyltransferase ZDHHC13	Generic enzyme



**Table 4.0 ANOVA Cluster 2**

#	Gene Symbol	Protein name	Object type
1	<u>ACSL6</u>	Long-chain-fatty-acid--CoA ligase 6	Generic enzyme
2	<u>APOL1</u>	Apolipoprotein L1	Receptor ligand
3	<u>BDNF</u>	Brain-derived neurotrophic factor	Receptor ligand
4	<u>BTN3A2</u>	Butyrophilin subfamily 3 member A2	Generic protein
5	<u>BTN3A3</u>	Butyrophilin subfamily 3 member A3	Generic protein
6	<u>C11orf82</u>	Nitric oxide-inducible gene protein	Generic protein
7	<u>C12orf4</u>	Uncharacterized protein C12orf4	Generic protein
8	<u>C1S</u>	Basic proline-rich peptide IB-1	Generic binding protein
		Complement C1s subcomponent	Generic protease
9	<u>C3</u>	Complement C3	Generic binding protein
10	<u>C3orf14</u>	Uncharacterized protein C3orf14	Generic protein
11	<u>CASP1</u>	Caspase-1	Generic protease
12	<u>CCNA2</u>	Cyclin-A2	Generic binding protein
13	<u>CCNB2</u>	G2/mitotic-specific cyclin-B2	Generic binding protein
14	<u>CD74</u>	HLA class II histocompatibility antigen gamma chain	Generic receptor
15	<u>CDC25B</u>	M-phase inducer phosphatase 2	Protein phosphatase
16	<u>CDO1</u>	Cysteine dioxygenase type 1	Generic enzyme
17	<u>COQ3</u>	Hexaprenyldihydroxybenzoate methyltransferase, mitochondrial	Generic enzyme
18	<u>COX6A1</u>	Cytochrome c oxidase subunit 6A1, mitochondrial	Generic enzyme
19	<u>CPXM1</u>	Probable carboxypeptidase X1	Generic protease
20	<u>CTSS</u>	Cathepsin S	Generic protease
21	<u>CXCL11</u>	C-X-C motif chemokine 11	Receptor ligand
22	<u>DKC1</u>	H/ACA ribonucleoprotein complex subunit 4	Generic enzyme
23	<u>DPY19L2</u>	Protein dpy-19 homolog 2	Generic protein
24	<u>DPY19L2P1</u>	Protein dpy-19 homolog 2-like 1	Generic protein
25	<u>DPY19L2P2</u>	dpy-19-like 2 pseudogene 2 (C. elegans)	Generic protein
26	<u>DPY19L2P4</u>	dpy-19-like 2 pseudogene 4 (C. elegans)	Generic protein
27	<u>ERP27</u>	Endoplasmic reticulum resident protein ERp27	Generic protein
28	<u>FAM163A</u>	UPF0417 protein FAM163A	Generic binding protein
29	<u>FARSB</u>	Phenylalanyl-tRNA synthetase beta chain	Generic enzyme
30	<u>FAS</u>	Tumor necrosis factor receptor superfamily member 6	Generic receptor
31	<u>FKBP5</u>	FK506-binding protein 5	Generic binding protein
32	<u>GBP1</u>	Interferon-induced guanylate-binding protein 1	Generic binding protein

33	<u>GBP3</u>	Guanylate-binding protein 3	Generic enzyme
34	<u>GGH</u>	Gamma-glutamyl hydrolase	Generic protease
35	<u>GLIPR1</u>	Glioma pathogenesis-related protein 1	Generic protein
36	<u>GNAS</u>	Guanine nucleotide-binding protein G(s) subunit alpha isoforms short	G-alpha
		Neuroendocrine secretory protein 55	Generic enzyme
		Protein ALEX	G-alpha
		Guanine nucleotide-binding protein G(s) subunit alpha isoforms XLas	G-alpha
37	<u>GPBR</u>	G-protein coupled estrogen receptor 1	GPCR
38	<u>HIST1H2BK</u>	Histone H2B type 1-K	Generic binding protein
39	<u>IFI16</u>	Gamma-interferon-inducible protein Ifi-16	Generic binding protein
40	<u>IFI30</u>	Gamma-interferon-inducible lysosomal thiol reductase	Generic enzyme
41	<u>IFITM2</u>	Interferon-induced transmembrane protein 2	Generic binding protein
42	<u>IGFBP5</u>	Insulin-like growth factor-binding protein 5	Generic binding protein
43	<u>IL10RB</u>	Interleukin-10 receptor beta chain	Generic receptor
44	<u>IL1R2</u>	Interleukin-1 receptor type II	Generic receptor
45	<u>IRF1</u>	Interferon regulatory factor 1	Transcription factor
46	<u>KIF16B</u>	Kinesin-like protein KIF16B	Generic binding protein
47	<u>KLF4</u>	Krueppel-like factor 4	Transcription factor
48	<u>KYNU</u>	Kynureninase	Generic enzyme
49	<u>LGALS1</u>	Galectin-1	Receptor ligand
50	<u>LOC100131642</u>		
51	<u>LOC653879</u>		
52	<u>ME1</u>	NADP-dependent malic enzyme	Generic enzyme
53	<u>MKI67</u>	Antigen KI-67	Generic binding protein
54	<u>MLKL</u>	Mixed lineage kinase domain-like protein	Protein kinase
55	<u>NCOA7</u>	Nuclear receptor coactivator 7	Generic binding protein
56	<u>NID2</u>	Nidogen-2	Generic binding protein
57	<u>NMI</u>	N-myc-interactor	Generic binding protein
58	<u>NQO1</u>	NAD(P)H dehydrogenase [quinone] 1	Generic enzyme
59	<u>PCDH19</u>	Protocadherin-19	Generic binding protein
60	<u>PIK3R2</u>	Phosphatidylinositol 3-kinase regulatory subunit beta	Generic binding protein
61	<u>PML</u>	Probable transcription factor PML	Transcription factor
62	<u>PNO1</u>	RNA-binding protein PNO1	Generic binding protein
63	<u>POP5</u>	Ribonuclease P/MRP protein subunit POP5	Generic enzyme
64	<u>PSMB9</u>	Proteasome subunit beta type-9	Generic protease
65	<u>PSMC5</u>	26S protease regulatory subunit 8	Generic enzyme

66	<u>PSME2</u>	Proteasome activator complex subunit 2	Generic binding protein
67	<u>PTN</u>	Pleiotrophin	Receptor ligand
68	<u>RAP1B</u>	Ras-related protein Rap-1b	RAS superfamily
69	<u>RARRES3</u>	Retinoic acid receptor responder protein 3	Generic protein
70	<u>RNF5</u>	E3 ubiquitin-protein ligase RNF5	Generic enzyme
71	<u>RPLP0P1</u>		
72	<u>SAMD9L</u>	Sterile alpha motif domain-containing protein 9-like	Generic protein
73	<u>SERPING1</u>	Plasma protease C1 inhibitor	Generic binding protein
74	<u>SFRS2</u>	Splicing factor, arginine/serine-rich 2	Generic binding protein
75	<u>SLC15A3</u>	Solute carrier family 15 member 3	Transporter
76	<u>SLC25A12</u>	Calcium-binding mitochondrial carrier protein Aralar1	Transporter
77	<u>SNRPF</u>	Small nuclear ribonucleoprotein F	Generic binding protein
78	<u>SP140L</u>	Nuclear body protein SP140-like protein	Generic binding protein
79	<u>STAMBPL1</u>	AMSH-like protease	Metalloprotease
80	<u>STRN3</u>	Striatin-3	Generic binding protein
81	<u>TAP1</u>	Antigen peptide transporter 1	Transporter
82	<u>TAP2</u>	Antigen peptide transporter 2	Transporter
83	<u>TARP</u>		
84	<u>TFRC</u>	Transferrin receptor protein 1	Generic receptor
85	<u>TLR3</u>	Toll-like receptor 3	Generic receptor
86	<u>TMEM60</u>	Transmembrane protein 60	Generic protein
87	<u>TMEM97</u>	Transmembrane protein 97	Generic protein
88	<u>TRGV3</u>	T-cell receptor gamma chain V region PT-gamma-1/2	Generic receptor
89	<u>TRGV5</u>	T cell receptor gamma variable 5	Generic receptor
90	<u>TRIM22</u>	E3 ubiquitin-protein ligase TRIM22	Transcription factor
91	<u>UBFD1</u>	Ubiquitin domain-containing protein UBFD1	Generic protein
92	<u>UCHL1</u>	Ubiquitin carboxyl-terminal hydrolase isozyme L1	Generic protease
93	<u>VEGFC</u>	Vascular endothelial growth factor C	Receptor ligand
94	<u>XPR1</u>	Xenotropic and polytropic retrovirus receptor 1	GPCR
95	<u>ZNF236</u>	Zinc finger protein 236	Transcription factor
96	<u>hCG_1757335</u>	Ras-related protein Rap-1b-like protein	RAS superfamily

**Table 5.0 ANOVA Cluster 3**

#	Gene Symbol	Protein name	Object type
1	<u>ACIN1</u>	Apoptotic chromatin condensation inducer in the nucleus	Generic binding protein
2	<u>AGRN</u>	Agrin	Generic binding protein
3	<u>AKAP2</u>	A-kinase anchor protein 2	Generic binding protein
4	<u>AKT1</u>	RAC-alpha serine/threonine-protein kinase	Protein kinase
5	<u>ANXA1</u>	Annexin A1	Generic binding protein
6	<u>BMP1</u>	Bone morphogenetic protein 1	Metalloprotease
7	<u>BRMS1</u>	Breast cancer metastasis-suppressor 1	Generic binding protein
8	<u>C14orf156</u>	SRA stem-loop-interacting RNA-binding protein, mitochondrial	Generic binding protein
9	<u>C17orf71</u>	Protein SMG8	Generic binding protein
10	<u>C19orf28</u>	Uncharacterized MFS-type transporter C19orf28	Generic protein
11	<u>C1orf43</u>	Uncharacterized protein C1orf43	Generic protein
12	<u>C2orf29</u>	Uncharacterized protein C2orf29	Generic protein
13	<u>CALM1</u>	Calmodulin	Generic binding protein
14	<u>CALM2</u>	Calmodulin	Generic binding protein
15	<u>CALM3</u>	Calmodulin	Generic binding protein
16	<u>CBR3</u>	Carbonyl reductase [NADPH] 3	Generic enzyme
17	<u>CC2D1A</u>	Coiled-coil and C2 domain-containing protein 1A	Generic binding protein
18	<u>CCDC85C</u>	Coiled-coil domain-containing protein 85C	Generic protein
19	<u>CCND3</u>	G1/S-specific cyclin-D3	Generic binding protein
20	<u>CDC6</u>	Cell division control protein 6 homolog	Generic binding protein
21	<u>CDS2</u>	Phosphatidate cytidyltransferase 2	Generic enzyme
22	<u>CHCHD3</u>	Coiled-coil-helix-coiled-coil-helix domain-containing protein 3, mitochondrial	Generic binding protein
23	<u>CNN3</u>	Calponin-3	Generic binding protein
24	<u>CSNK1E</u>	Casein kinase I isoform epsilon	Protein kinase
25	<u>CSRP2BP</u>	Cysteine-rich protein 2-binding protein	Generic binding protein
26	<u>DDX58</u>	Probable ATP-dependent RNA helicase DDX58	Generic enzyme
27	<u>DTX3L</u>	E3 ubiquitin-protein ligase DTX3L	Generic binding protein
28	<u>EIF3B</u>	Eukaryotic translation initiation factor 3 subunit B	Generic binding protein
29	<u>EPSTI1</u>	Epithelial-stromal interaction protein 1	Generic protein
30	<u>EXOC7</u>	Exocyst complex component 7	Generic protein
31	<u>FAM100A</u>	Protein FAM100A	Generic protein
32	<u>FAM125A</u>	Multivesicular body subunit 12A	Generic binding protein

33	<u>FAM38A</u>	Protein FAM38A	Generic protein
34	<u>FAM53C</u>	Protein FAM53C	Generic protein
35	<u>FBXO4</u>	F-box only protein 4	Generic binding protein
36	<u>FILIP1L</u>	Filamin A-interacting protein 1-like	Generic protein
37	<u>FKRP</u>	Fukutin-related protein	Generic enzyme
38	<u>GDPD3</u>	Glycerophosphodiester phosphodiesterase domain-containing protein 3	Generic enzyme
39	<u>HAS3</u>	Hyaluronan synthase 3	Generic enzyme
40	<u>HEXIM1</u>	Protein HEXIM1	Generic binding protein
41	<u>HNRNPAB</u>	Heterogeneous nuclear ribonucleoprotein A/B	Generic binding protein
42	<u>HNRNPM</u>	Heterogeneous nuclear ribonucleoprotein M	Generic binding protein
43	<u>HSPA9</u>	Stress-70 protein, mitochondrial	Generic binding protein
44	<u>IFI44L</u>	Interferon-induced protein 44-like	Generic protease
45	<u>IFIT3</u>	Interferon-induced protein with tetratricopeptide repeats 3	Generic binding protein
46	<u>IFITM1</u>	Interferon-induced transmembrane protein 1	Generic binding protein
47	<u>IFITM3</u>	Interferon-induced transmembrane protein 3	Generic protein
48	<u>ILF3</u>	Interleukin enhancer-binding factor 3	Transcription factor
49	<u>INO80C</u>	INO80 complex subunit C	Generic protein
50	<u>IRF1</u>	Interferon regulatory factor 1	Transcription factor
51	<u>KDM3B</u>	Lysine-specific demethylase 3B	Generic enzyme
52	<u>KIF23</u>	Kinesin-like protein KIF23	Generic binding protein
53	<u>KLF13</u>	Krueppel-like factor 13	Transcription factor
54	<u>KLF6</u>	Krueppel-like factor 6	Transcription factor
55	<u>KPNA2</u>	Importin subunit alpha-2	Transporter
56	<u>LAMP3</u>	Lysosome-associated membrane glycoprotein 3	Generic protein
57	<u>LAP3</u>	Cytosol aminopeptidase	Generic protease
58	<u>LARP1</u>	La-related protein 1	Generic binding protein
59	<u>LMF2</u>	Lipase maturation factor 2	Generic protein
60	<u>LOC388796</u>		
61	<u>LOC652147</u>		
62	<u>LOC727996</u>		
63	<u>LOC728216</u>		
64	<u>LOC728438</u>		
65	<u>LSG1</u>	Large subunit GTPase 1 homolog	Generic binding protein
66	<u>LSM12</u>	Protein LSM12 homolog	Generic binding protein
67	<u>MARVELD1</u>	Putative MARVEL domain-containing protein 1	Generic protein
68	<u>MAT2A</u>	S-adenosylmethionine synthetase isoform type-2	Generic enzyme
69	<u>MCM3</u>	DNA replication licensing factor MCM3	Generic binding protein
70	<u>MED22</u>	Mediator of RNA polymerase II transcription subunit 22	Generic binding protein

71	<u>MGST2</u>	Microsomal glutathione S-transferase 2	Generic enzyme
72	<u>MICALL1</u>	MICAL-like protein 1	Generic binding protein
73	<u>MRPL4</u>	39S ribosomal protein L4, mitochondrial	Generic binding protein
74	<u>MSL1</u>	Male-specific lethal 1 homolog	Generic binding protein
75	<u>MUC1</u>	Mucin-1	Generic binding protein
76	<u>MX1</u>	Interferon-induced GTP-binding protein Mx1	Generic enzyme
77	<u>MX2</u>	Interferon-induced GTP-binding protein Mx2	Generic enzyme
78	<u>NACC1</u>	Nucleus accumbens-associated protein 1	Generic binding protein
79	<u>NLRC5</u>	Protein NLRC5	Generic binding protein
80	<u>OAS1</u>	2'-5'-oligoadenylate synthetase 1	Generic enzyme
81	<u>OAS2</u>	2'-5'-oligoadenylate synthetase 2	Generic enzyme
82	<u>OAS3</u>	2'-5'-oligoadenylate synthetase 3	Generic enzyme
83	<u>OR7E104P</u>		
84	<u>PACS2</u>	Phosphofurin acidic cluster sorting protein 2	Generic protein
85	<u>PALM2</u>	Paralemmin-2	Generic kinase
86	<u>PALM2-AKAP2</u>	OTTHUMP00000063892	Generic binding protein
87	<u>PARP1</u>	Poly [ADP-ribose] polymerase 1	Generic enzyme
88	<u>PARP9</u>	Poly [ADP-ribose] polymerase 9	Generic enzyme
89	<u>PCGF5</u>	Polycomb group RING finger protein 5	Generic binding protein
90	<u>PCNT</u>	Pericentrin	Generic binding protein
91	<u>PCNXL3</u>	Pecanex-like protein 3	Generic protein
92	<u>PDCD11</u>	Protein RRP5 homolog	Generic binding protein
93	<u>PFN1</u>	Profilin-1	Generic binding protein
94	<u>PIGU</u>	Phosphatidylinositol glycan anchor biosynthesis class U protein	Generic binding protein
95	<u>PLK2</u>	Serine/threonine-protein kinase PLK2	Protein kinase
96	<u>PPIL5</u>	Peptidylprolyl isomerase-like 5	Generic enzyme
97	<u>PPP2CA</u>	Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform	Protein phosphatase
98	<u>RAN</u>	GTP-binding nuclear protein Ran	RAS superfamily
99	<u>RBBP4</u>	Histone-binding protein RBBP4	Generic enzyme
100	<u>RHOB</u>	Rho-related GTP-binding protein RhoB	RAS superfamily
101	<u>ROR1</u>	Tyrosine-protein kinase transmembrane receptor ROR1	Receptor with enzyme activity
102	<u>RYR1</u>	Ryanodine receptor 1	Ligand-gated ion-channel
103	<u>SAMD9</u>	Sterile alpha motif domain-containing protein 9	Generic protein
104	<u>SELI</u>	Ethanolaminephosphotransferase 1	Generic enzyme
105	<u>SEPHS1</u>	Selenide, water dikinase 1	Generic kinase
106	<u>SFRS1</u>	Splicing factor, arginine/serine-rich 1	Generic binding protein

107	<u>SFRS3</u>	Splicing factor, arginine/serine-rich 3	Generic binding protein
108	<u>SGPL1</u>	Sphingosine-1-phosphate lyase 1	Generic enzyme
109	<u>SLC25A20</u>	Mitochondrial carnitine/acylcarnitine carrier protein	Transporter
110	<u>SLC38A7</u>	Putative sodium-coupled neutral amino acid transporter 7	Generic binding protein
111	<u>SMARCA4</u>	Probable global transcription activator SNF2L4	Generic enzyme
112	<u>SP110</u>	Sp110 nuclear body protein	Generic binding protein
113	<u>SRRT</u>	Serrate RNA effector molecule homolog	Generic binding protein
114	<u>STAT1</u>	Signal transducer and activator of transcription 1-alpha/beta	Transcription factor
115	<u>STEAP3</u>	Metalloreductase STEAP3	Generic enzyme
116	<u>SYNCRIP</u>	Heterogeneous nuclear ribonucleoprotein Q	Generic binding protein
117	<u>TACC3</u>	Transforming acidic coiled-coil-containing protein 3	Generic binding protein
118	<u>TAGLN2</u>	Transgelin-2	Generic protein
119	<u>TEAD4</u>	Transcriptional enhancer factor TEF-3	Transcription factor
120	<u>TEX261</u>	Protein TEX261	Generic protein
121	<u>TGOLN2</u>	Trans-Golgi network integral membrane protein 2	Generic binding protein
122	<u>THBS1</u>	Thrombospondin-1	Receptor ligand
123	<u>TIMM22</u>	Mitochondrial import inner membrane translocase subunit Tim22	Transporter
124	<u>TK1</u>	Thymidine kinase, cytosolic	Generic kinase
125	<u>TMEM140</u>	Transmembrane protein 140	Generic protein
126	<u>TMEM173</u>	Transmembrane protein 173	Generic protein
127	<u>TMEM201</u>	Transmembrane protein 201	Generic protein
128	<u>TMPO</u>	Lamina-associated polypeptide 2, isoform alpha	Generic binding protein
		Lamina-associated polypeptide 2, isoforms beta/gamma	Generic binding protein
129	<u>TPX2</u>	Targeting protein for Xklp2	Generic binding protein
130	<u>TRAFD1</u>	TRAF-type zinc finger domain-containing protein 1	Generic binding protein
131	<u>TRIM21</u>	52 kDa Ro protein	Generic binding protein
132	<u>TRIM26</u>	Tripartite motif-containing protein 26	Generic binding protein
133	<u>TRIM69</u>	Tripartite motif-containing protein 69	Generic binding protein
134	<u>TRIP13</u>	Thyroid receptor-interacting protein 13	Generic binding protein
135	<u>UBE3C</u>	Ubiquitin-protein ligase E3C	Generic enzyme
136	<u>UHRF1</u>	E3 ubiquitin-protein ligase UHRF1	Generic enzyme
137	<u>UMPS</u>	Uridine 5'-monophosphate synthase	Generic enzyme

138	<u>USP18</u>	Ubl carboxyl-terminal hydrolase 18	Generic protease
139	<u>WDR77</u>	Methylosome protein 50	Generic binding protein
140	<u>XAF1</u>	XIAP-associated factor 1	Generic binding protein
141	<u>YWHAB</u>	14-3-3 protein beta/alpha	Generic binding protein
142	<u>ZWINT</u>	ZW10 interactor	Generic binding protein

**Table 6.0 ANOVA Cluster 4**

#	Gene Symbol	Protein name	Object type
1	<u>ADAMTS1</u>	A disintegrin and metalloproteinase with thrombospondin motifs 1	Metalloprotease
2	<u>AGPAT9</u>	Glycerol-3-phosphate acyltransferase 3	Generic enzyme
3	<u>AHNAK2</u>	Protein AHNAK2	Generic binding protein
4	<u>AKAP12</u>	A-kinase anchor protein 12	Generic binding protein
5	<u>AKAP13</u>	A-kinase anchor protein 13	Regulators (GDI, GAP, GEF etc.)
6	<u>ALOX5</u>	Arachidonate 5-lipoxygenase	Generic enzyme
7	<u>ANKRD37</u>	Ankyrin repeat domain-containing protein 37	Generic binding protein
8	<u>AP1S1</u>	AP-1 complex subunit sigma-1A	Transporter
9	<u>APBB2</u>	Amyloid beta A4 precursor protein-binding family B member 2	Generic binding protein
10	<u>ARG2</u>	Arginase-2, mitochondrial	Generic enzyme
11	<u>ARHGAP22</u>	Rho GTPase-activating protein 22	Regulators (GDI, GAP, GEF etc.)
12	<u>ARHGAP5</u>	Rho GTPase-activating protein 5	Regulators (GDI, GAP, GEF etc.)
13	<u>ARL1</u>	ADP-ribosylation factor-like protein 1	RAS superfamily
14	<u>ARL4C</u>	ADP-ribosylation factor-like protein 4C	RAS superfamily
15	<u>ATP1B1</u>	Sodium/potassium-transporting ATPase subunit beta-1	Generic binding protein
16	<u>BHLHE40</u>	Class E basic helix-loop-helix protein 40	Transcription factor
17	<u>BTBD7</u>	BTB/POZ domain-containing protein 7	Generic binding protein
18	<u>C12orf76</u>	Uncharacterized protein C12orf76	Generic protein
19	<u>C13orf36</u>	Uncharacterized protein C13orf36	Generic protein
20	<u>C18orf19</u>	Uncharacterized protein C18orf19	Generic protein
21	<u>C4orf47</u>	UPF0602 protein C4orf47	Generic protein
22	<u>C4orf49</u>	Uncharacterized protein C4orf49	Generic protein
23	<u>C7orf68</u>	Hypoxia-inducible gene 2 protein	Generic binding protein
24	<u>CAMK2D</u>	Calcium/calmodulin-dependent protein kinase type II delta chain	Protein kinase
25	<u>CBFA2T2</u>	Protein CBFA2T2	Transcription factor
26	<u>CCNG1</u>	Cyclin-G1	Generic binding protein



27	<u>CD44</u>	CD44 antigen	Generic receptor
28	<u>CD46</u>	Membrane cofactor protein	Generic receptor
29	<u>CD55</u>	Complement decay-accelerating factor	Generic binding protein
30	<u>CEP57</u>	Centrosomal protein of 57 kDa	Generic binding protein
31	<u>CHPT1</u>	Cholinephosphotransferase 1	Generic enzyme
32	<u>CLEC11A</u>	C-type lectin domain family 11 member A	Receptor ligand
33	<u>CPD</u>	Carboxypeptidase D	Generic protease
34	<u>CTBS</u>	Di-N-acetylchitobiase	Generic enzyme
35	<u>CTSB</u>	Cathepsin B	Generic protease
36	<u>CTTN</u>	Src substrate cortactin	Generic binding protein
37	<u>CXCR4</u>	C-X-C chemokine receptor type 4	GPCR
38	<u>DCBLD2</u>	Discoidin, CUB and LCCL domain-containing protein 2	Generic binding protein
39	<u>DCUN1D4</u>	DCN1-like protein 4	Generic protein
40	<u>DPY30</u>	Protein dpy-30 homolog	Generic binding protein
41	<u>DRAM2</u>	DNA damage-regulated autophagy modulator protein 2	Generic protein
42	<u>EGFR</u>	Epidermal growth factor receptor	Receptor with enzyme activity
43	<u>EHF</u>	ETS homologous factor	Transcription factor
44	<u>EIF3D</u>	Eukaryotic translation initiation factor 3 subunit D	Generic binding protein
45	<u>EIF3E</u>	Eukaryotic translation initiation factor 3 subunit E	Generic binding protein
46	<u>ELTD1</u>	EGF, latrophilin and seven transmembrane domain-containing protein 1	GPCR
47	<u>EPRS</u>	Bifunctional aminoacyl-tRNA synthetase	Generic enzyme
48	<u>ERN1</u>	Serine/threonine-protein kinase/endoribonuclease IRE1	Protein kinase
49	<u>ERO1L</u>	ERO1-like protein alpha	Generic enzyme
50	<u>ERP44</u>	Endoplasmic reticulum resident protein ERp44	Generic binding protein
51	<u>ETNK1</u>	Ethanolamine kinase 1	Generic kinase
52	<u>ETV3</u>	ETS translocation variant 3	Transcription factor
53	<u>FAM110C</u>	Protein FAM110C	Generic protein
54	<u>FAM162A</u>	UPF0389 protein FAM162A	Generic protein
55	<u>FAM167A</u>	UPF0484 protein FAM167A	Generic protein
56	<u>FAM45A</u>	Protein FAM45A	Generic protein
57	<u>FAM45B</u>	Protein FAM45B	Generic protein
58	<u>FIBCD1</u>	Fibrinogen C domain-containing protein 1	Generic binding protein
59	<u>FNTA</u>	Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha	Generic enzyme
60	<u>FOXL1</u>	Forkhead box protein L1	Transcription factor

61	<u>FXR1</u>	Fragile X mental retardation syndrome-related protein 1	Generic binding protein
62	<u>FZD10</u>	Frizzled-10	GPCR
63	<u>G0S2</u>	Putative lymphocyte G0/G1 switch protein 2	Generic protein
64	<u>GALNT5</u>	Polypeptide N-acetylgalactosaminyltransferase 5	Generic enzyme
65	<u>GALNTL2</u>	Polypeptide N-acetylgalactosaminyltransferase-like protein 2	Generic enzyme
66	<u>GJC1</u>	Gap junction gamma-1 protein	Generic channel
67	<u>GLS</u>	Glutaminase kidney isoform, mitochondrial	Generic enzyme
68	<u>GLT8D1</u>	Glycosyltransferase 8 domain-containing protein 1	Generic enzyme
69	<u>GOLGB1</u>	Golgin subfamily B member 1	Generic binding protein
70	<u>GOLT1A</u>	Vesicle transport protein GOT1A	Generic binding protein
71	<u>GPR155</u>	Integral membrane protein GPR155	Generic receptor
72	<u>GYS2</u>	Glycogen [starch] synthase, liver	Generic enzyme
73	<u>HLTF</u>	Helicase-like transcription factor	Generic binding protein
74	<u>HOMER1</u>	Homer protein homolog 1	Generic binding protein
75	<u>HSP90B1</u>	Endoplasmic	Generic binding protein
76	<u>IGFBP3</u>	Insulin-like growth factor-binding protein 3	Generic binding protein
77	<u>ING2</u>	Inhibitor of growth protein 2	Generic binding protein
78	<u>INSC</u>	Protein inscuteable homolog	Generic binding protein
79	<u>INSIG2</u>	Insulin-induced gene 2 protein	Generic binding protein
80	<u>INTS6</u>	Integrator complex subunit 6	Generic enzyme
81	<u>IRF6</u>	Interferon regulatory factor 6	Transcription factor
82	<u>IVNS1ABP</u>	Influenza virus NS1A-binding protein	Generic binding protein
83	<u>JARID2</u>	Protein Jumonji	Generic binding protein
84	<u>KDM3A</u>	Lysine-specific demethylase 3A	Generic enzyme
85	<u>KDM5B</u>	Lysine-specific demethylase 5B	Transcription factor
86	<u>KIAA0430</u>	Limkain-b1	Generic binding protein
87	<u>KIAA1143</u>	Uncharacterized protein KIAA1143	Generic protein
88	<u>KIAA1244</u>	Brefeldin A-inhibited guanine nucleotide-exchange protein 3	Regulators (GDI, GAP, GEF etc.)
89	<u>KISS1R</u>	KiSS-1 receptor	GPCR
90	<u>KLHL24</u>	Kelch-like protein 24	Generic binding protein
91	<u>LAMB3</u>	Laminin subunit beta-3	Receptor ligand
92	<u>LAMC2</u>	Laminin subunit gamma-2	Receptor ligand
93	<u>LAPTM4A</u>	Lysosomal-associated transmembrane protein 4A	Generic protein
94	<u>LEMD1</u>	LEM domain-containing protein 1	Generic protein
95	<u>LEPR</u>	Leptin receptor	Generic receptor

96	<u>LGR5</u>	Leucine-rich repeat-containing G-protein coupled receptor 5	GPCR
97	<u>LIMA1</u>	LIM domain and actin-binding protein 1	Generic binding protein
98	<u>LIPG</u>	Endothelial lipase	Generic phospholipase
99	<u>LMAN1</u>	Protein ERGIC-53	Generic binding protein
100	<u>LOC149832</u>		
101	<u>LOC387763</u>	Protein Ag2 homolog	Generic protein
102	<u>LOC554202</u>	Putative uncharacterized protein LOC554202	Generic protein
103	<u>LOC729082</u>		
104	<u>LOC729222</u>		
105	<u>LOX</u>	Protein-lysine 6-oxidase	Generic enzyme
106	<u>LRRC16A</u>	Leucine-rich repeat-containing protein 16A	Generic binding protein
107	<u>LRRFIP1</u>	Leucine-rich repeat flightless-interacting protein 1	Transcription factor
108	<u>LTA4H</u>	Leukotriene A-4 hydrolase	Generic enzyme
109	<u>MAN2A1</u>	Alpha-mannosidase 2	Generic enzyme
110	<u>MAP2</u>	Microtubule-associated protein 2	Generic binding protein
111	<u>MAP4K4</u>	Mitogen-activated protein kinase kinase kinase kinase 4	Protein kinase
112	<u>MAST4</u>	Microtubule-associated serine/threonine-protein kinase 4	Protein kinase
113	<u>MBOAT2</u>	Lysophospholipid acyltransferase 2	Generic enzyme
114	<u>METTL5</u>	Methyltransferase-like protein 5	Generic enzyme
115	<u>MIB1</u>	E3 ubiquitin-protein ligase MIB1	Generic enzyme
116	<u>MPZL2</u>	Myelin protein zero-like protein 2	Generic binding protein
117	<u>MREG</u>	Melanoregulin	Generic protein
118	<u>MRI1</u>	Methylthioribose-1-phosphate isomerase	Generic binding protein
119	<u>MTSS1</u>	Metastasis suppressor protein 1	Generic binding protein
120	<u>MXI1</u>	MAX-interacting protein 1	Transcription factor
121	<u>NCRNA00201</u>		
122	<u>NCRNA00219</u>	Putative uncharacterized protein TIGA1	Generic protein
123	<u>NDRG1</u>	Protein NDRG1	Generic binding protein
124	<u>NFKBIZ</u>	NF-kappa-B inhibitor zeta	Generic binding protein
125	<u>NGLY1</u>	Peptide-N(4)-(N-acetyl-beta-glucosaminy)asparagine amidase	Generic enzyme
126	<u>NOTCH2NL</u>	Notch homolog 2 N-terminal-like protein	Generic binding protein
127	<u>NRIP3</u>	Nuclear receptor-interacting protein 3	Generic binding protein
128	<u>NT5E</u>	5'-nucleotidase	Generic phosphatase
129	<u>OPN3</u>	Opsin-3	GPCR
130	<u>PAM</u>	Peptidyl-glycine alpha-amidating monooxygenase	Generic enzyme

131	<u>PDGFC</u>	Platelet-derived growth factor C	Receptor ligand
132	<u>PECI</u>	Peroxisomal 3,2-trans-enoyl-CoA isomerase	Generic enzyme
133	<u>PHF20</u>	PHD finger protein 20	Generic binding protein
134	<u>PHLDA1</u>	Pleckstrin homology-like domain family A member 1	Generic binding protein
135	<u>PIK3CB</u>	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta isoform	Lipid kinase
136	<u>PLAT</u>	Tissue-type plasminogen activator	Generic protease
137	<u>PLCG2</u>	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma-2	Generic phospholipase
138	<u>PPFIBP1</u>	Liprin-beta-1	Generic binding protein
139	<u>PPP1R3B</u>	Protein phosphatase 1 regulatory subunit 3B	Protein phosphatase
140	<u>PRKRA</u>	Interferon-inducible double stranded RNA-dependent protein kinase activator A	Generic binding protein
141	<u>PRUNE2</u>	Protein prune homolog 2	Generic protein
		BNIP2 motif-containing molecule at the C-terminal region 1	Generic protein
142	<u>PSG1</u>	Pregnancy-specific beta-1-glycoprotein 1	Generic protein
143	<u>PTEN</u>	Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase PTEN	Lipid phosphatase
144	<u>QPCT</u>	Glutamyl-peptide cyclotransferase	Generic enzyme
145	<u>RAB13</u>	Ras-related protein Rab-13	RAS superfamily
146	<u>RAP2B</u>	Ras-related protein Rap-2b	RAS superfamily
147	<u>RBM5</u>	RNA-binding protein 5	Generic binding protein
148	<u>RBM6</u>	RNA-binding protein 6	Generic binding protein
149	<u>RBPJ</u>	Recombining binding protein suppressor of hairless	Transcription factor
150	<u>RGS20</u>	Regulator of G-protein signaling 20	Regulators (GDI, GAP, GEF etc.)
151	<u>RNASE4</u>	Ribonuclease 4	Generic enzyme
152	<u>RPL22</u>	60S ribosomal protein L22	Generic binding protein
153	<u>RPL31</u>	60S ribosomal protein L31	Generic binding protein
154	<u>RPL31P10</u>		
155	<u>RPL31P17</u>		
156	<u>RPL31P18</u>		
157	<u>RPL31P39</u>		
158	<u>RPL31P4</u>	similar to ribosomal protein L31	Generic protein
159	<u>RPL5</u>	60S ribosomal protein L5	Generic binding protein
160	<u>RRAS2</u>	Ras-related protein R-Ras2	RAS superfamily
161	<u>SAT1</u>	Diamine acetyltransferase 1	Generic enzyme

162	<u>SATB1</u>	DNA-binding protein SATB1	Transcription factor
163	<u>SCAMP1</u>	Secretory carrier-associated membrane protein 1	Generic binding protein
164	<u>SCD</u>	Acyl-CoA desaturase	Generic enzyme
		Putative uncharacterized protein PRO1933	Generic protein
165	<u>SDC2</u>	Syndecan-2	Generic receptor
166	<u>SEL1L</u>	Protein sel-1 homolog 1	Generic binding protein
167	<u>SERTAD2</u>	SERTA domain-containing protein 2	Generic binding protein
168	<u>SH3YL1</u>	SH3 domain-containing YSC84-like protein 1	Generic binding protein
169	<u>SLAMF7</u>	SLAM family member 7	Generic receptor
170	<u>SLC2A14</u>	Solute carrier family 2, facilitated glucose transporter member 14	Transporter
171	<u>SLC2A3</u>	Solute carrier family 2, facilitated glucose transporter member 3	Transporter
172	<u>SLC5A3</u>	Sodium/myo-inositol cotransporter	Transporter
173	<u>SNX21</u>	Sorting nexin-21	Generic binding protein
174	<u>SPAG4</u>	Sperm-associated antigen 4 protein	Generic binding protein
175	<u>SPIRE1</u>	Protein spire homolog 1	Generic binding protein
176	<u>SPON1</u>	Spondin-1	Generic binding protein
177	<u>SPRY2</u>	Protein sprouty homolog 2	Generic binding protein
178	<u>SRGAP2P1</u>	SLIT-ROBO Rho GTPase activating protein 2 pseudogene 1	Generic protein
179	<u>SSTR1</u>	Somatostatin receptor type 1	GPCR
180	<u>TAF1D</u>	TATA box-binding protein-associated factor RNA polymerase I subunit D	Generic binding protein
181	<u>TAF9B</u>	Transcription initiation factor TFIID subunit 9B	Generic binding protein
182	<u>TCEA3</u>	Transcription elongation factor A protein 3	Generic binding protein
183	<u>TFB1M</u>	Dimethyladenosine transferase 1, mitochondrial	Generic enzyme
184	<u>TG</u>	Thyroglobulin	Generic binding protein
185	<u>THBD</u>	Thrombomodulin	Generic receptor
186	<u>THSD4</u>	Thrombospondin type-1 domain-containing protein 4	Generic protease
187	<u>TLE1</u>	Transducin-like enhancer protein 1	Generic binding protein
188	<u>TMEM106B</u>	Transmembrane protein 106B	Generic protein
189	<u>TMEM117</u>	Transmembrane protein 117	Generic protein
190	<u>TMEM49</u>	Transmembrane protein 49	Generic protein
191	<u>TNFAIP3</u>	Tumor necrosis factor, alpha-induced protein 3	Generic binding protein
192	<u>TPP1</u>	Tripeptidyl-peptidase 1	Generic protease
193	<u>TRPM8</u>	Transient receptor potential cation channel subfamily M member 8	Generic channel

194	<u>TSHZ1</u>	Teashirt homolog 1	Generic binding protein
195	<u>TXNIP</u>	Thioredoxin-interacting protein	Generic binding protein
196	<u>UBE2B</u>	Ubiquitin-conjugating enzyme E2 B	Generic enzyme
197	<u>UBE2I</u>	SUMO-conjugating enzyme UBC9	Generic enzyme
198	<u>USO1</u>	General vesicular transport factor p115	Transporter
199	<u>USP10</u>	Ubiquitin carboxyl-terminal hydrolase 10	Generic protease
200	<u>VAMP1</u>	Vesicle-associated membrane protein 1	Generic binding protein
201	<u>VEGFA</u>	Vascular endothelial growth factor A	Receptor ligand
202	<u>VLDLR</u>	Very low-density lipoprotein receptor	Generic receptor
203	<u>WSB1</u>	WD repeat and SOCS box-containing protein 1	Generic binding protein
204	<u>XIST</u>		RNA
205	<u>XYLT2</u>	Xylosyltransferase 2	Generic enzyme
206	<u>YPEL4</u>	Protein yippee-like 4	Generic protein
207	<u>ZMYM2</u>	Zinc finger MYM-type protein 2	Generic binding protein
208	<u>ZNF185</u>	Zinc finger protein 185	Generic binding protein
209	<u>ZNF804A</u>	Zinc finger protein 804A	Generic binding protein

**Table 7.0 ANOVA Cluster 5**

#	Gene Symbol	Protein name	Object type
1	<u>ADM</u>	ADM	Receptor ligand
2	<u>AK3L1</u>	Adenylate kinase isoenzyme 4, mitochondrial	Generic kinase
3	<u>ANKMY2</u>	Ankyrin repeat and MYND domain-containing protein 2	Generic binding protein
4	<u>C18orf19</u>	Uncharacterized protein C18orf19	Generic protein
5	<u>CDH6</u>	Cadherin-6	Generic binding protein
6	<u>CLU</u>	Clusterin	Generic binding protein
7	<u>DEGS1</u>	Sphingolipid delta(4)-desaturase DES1	Generic enzyme
8	<u>EFEMP1</u>	EGF-containing fibulin-like extracellular matrix protein 1	Generic binding protein
9	<u>EIF2S2</u>	Eukaryotic translation initiation factor 2 subunit 2	Generic binding protein
10	<u>EPAS1</u>	Endothelial PAS domain-containing protein 1	Transcription factor
11	<u>ERRFI1</u>	ERBB receptor feedback inhibitor 1	Generic binding protein
12	<u>FAM13A</u>	Protein FAM13A	Generic protein
13	<u>FKBP11</u>	FK506-binding protein 11	Generic enzyme
14	<u>HEXB</u>	Beta-hexosaminidase subunit beta	Generic enzyme
15	<u>HSPA5</u>	78 kDa glucose-regulated protein	Generic binding protein
16	<u>HSPD1</u>	60 kDa heat shock protein, mitochondrial	Generic binding protein
17	<u>HSPH1</u>	Heat shock protein 105 kDa	Generic binding protein

18	<u>JMJD6</u>	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6	Generic enzyme
19	<u>LGALS8</u>	Galectin-8	Receptor ligand
20	<u>ME2</u>	NAD-dependent malic enzyme, mitochondrial	Generic enzyme
21	<u>MGC9913</u>	MGC9913	Generic protein
22	<u>MGST1</u>	Microsomal glutathione S-transferase 1	Generic enzyme
23	<u>MON2</u>	Protein MON2 homolog	Transporter
24	<u>P4HA2</u>	Prolyl 4-hydroxylase subunit alpha-2	Generic enzyme
25	<u>PAICS</u>	Multifunctional protein ADE2	Generic enzyme
26	<u>PCYOX1L</u>	Prenylcysteine oxidase-like	Generic enzyme
27	<u>PDIA5</u>	Protein disulfide-isomerase A5	Generic enzyme
28	<u>PDIA6</u>	Protein disulfide-isomerase A6	Generic enzyme
29	<u>PFKFB3</u>	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	Generic enzyme
30	<u>PGM1</u>	Phosphoglucomutase-1	Generic enzyme
31	<u>PPIL3</u>	Peptidyl-prolyl cis-trans isomerase-like 3	Generic enzyme
32	<u>SNHG1</u>		
33	<u>SORBS2</u>	Sorbin and SH3 domain-containing protein 2	Generic binding protein
34	<u>TFPI2</u>	Tissue factor pathway inhibitor 2	Generic binding protein

